CESifo WORKING PAPERS

6922
2018

Original Version: February 2018
This Version: February 2020

The International Organization of Production in the Regulatory Void

Philipp Herkenhoff, Sebastian Krautheim

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 6922*
Category 8: Trade Policy

The International Organization of Production in the Regulatory Void

Abstract

In recent decades, a large and increasing number of leading firms in a diverse set of industries have faced allegations of 'unethical' practices along their international value chains. In many cases this has triggered consumer boycotts and NGO campaigns, introducing a new link between upstream (un)ethical choices and downstream consumer demand. Does this feedback effect shape the international organization of production, as casual observation of NGO campaigns seems to suggest? In an explorative empirical investigation, we indeed find that - over and above the effects of established determinants of the integration vs outsourcing decision - high potential cost savings of 'unethical' production in an industry favor international outsourcing and most strongly so for sourcing from low-regulation countries. Motivated by these findings, we introduce a cost-saving 'unethical' technology, consumer boycotts, and advocacy NGOs into a standard property-rights model of the international organization of production. Contracts are incomplete, limiting a firm's control over both investment and (un)ethical technology choices of integrated as well as independent suppliers. We identify the unethical outsourcing incentive as a novel determinant of the international organization of production. It implies that high potential cost savings from 'unethical' production incentivize keeping the supplier at arm's length, rationalizing our empirical findings.

JEL-Codes: D230, F120, F180, F230, F610, L230, L310, Q530.

Keywords: international outsourcing, intrafirm trade, property-rights theory, environmental regulation, ethical production, consumer boycotts, NGOs.

Philipp Herkenhoff
Johannes Gutenberg-University
Mainz / Germany
philipp.herkenhoff@uni-mainz.de

*Sebastian Krautheim**
Chair of International Economics
University of Passau / Germany
sebastian.krautheim@alumni.eui.eu

*corresponding author

February 2020
We would like to thank Gani Aldashev, Pol Antràs, Gregory Corcos, Jonathan Eaton, Carsten Eckel, Gabriel Felbermayr, Bård Harstad, Wilhelm Kohler, Claudius Löhnert, Marcelo Olarreaga, Mathieu Parenti, Nina Pavcnik, Michael Pflüger, Gonzague Vannoorenberghe, Thierry Verdier, participants of the CESifo Area Conference on Global Economy, Munich; the 5th InsTED Workshop 'Advances in Institutions, Trade and Economic Development', Syracuse, NY; the 'Non-Profits, Governments, and Organizations Workshop' at the Graduate Institute, Geneva; the 19th International Economics Workshop, Göttingen; the 18th RIEF Doctoral Meetings in International Trade and International Finance, Munich; the 11th FIW-Research Conference 'International Economics', Vienna; the 4th Mainz Workshop on Foreign Direct Investment and Multinational Corporations; the 21st ETSG, Bern; as well as seminar participants at the Paris Trade Seminar (SciencesPo, INRA, Paris School of Economics, Université Paris 1), Université Catholique de Louvain, LMU Munich, University of Tübingen, University of Mainz, University of Würzburg, University of Hohenheim and University of Trier for fruitful discussions and comments. The Online Appendix is available upon request from the authors.

1 Introduction

The past three decades have been characterized by an unprecedented fragmentation and geographical dispersion of production. Value chains span all over the globe and even firms with a strong national branding have highly segmented international supply chains. This internationalization of production takes place in a context of substantial cross-country differences in environmental and labor regulation and international regulation of production is weak at best (Battaglini and Harstad, 2018). With heterogeneous national regulatory regimes, international production and sourcing decisions are taken in the *regulatory void* (Short, 2013).

With the expansion of international production, firms and their suppliers are more and more frequently accused of using the regulatory void to cut costs at the expense of the environment, local workers and future generations. There is a large number of examples in which allegations of 'unethical' or 'immoral' practices have led to consumer boycotts and NGO campaigns against global players in a diverse set of industries.[1] In addition to highlighting that upstream (un)ethical technology choices can have a negative feedback effect on downstream demand, these examples exhibit a striking pattern: in all cases, criticism of 'unethical' practices concerns actions by *independent* suppliers rather than subsidiaries of multinational firms.

Does the choice between 'ethical' and 'unethical' production interact with the international organization of production? Simply put: do firms delegate the 'dirty work' of globalization to independent suppliers?

As we are interested in the choices of firms in international value chains, the literature on the international organization of production based on the seminal contribution by Antràs (2003) provides the natural framework to address this question. It applies the property-rights theory of the firm by Grossman and Hart (1986) and Hart and Moore (1990) to the international organization of production. This literature has identified several key determinants of the international make-or-buy-decision both theoretically and empirically. To our knowledge, the regulatory void and consumer boycotts are not among them. To go beyond anecdotal evidence, we follow the empirical branch of this literature (e.g. Nunn and Trefler 2013, Antràs and Chor 2013, Antràs and Yeaple 2014, Antràs 2016) and provide an explorative empirical analysis of the issue in Section 2. Specifically, we investigate whether firms in industries with stronger incentives for 'unethical' production tend to keep their suppliers at arm's length especially when sourcing from low-regulation countries.

To do so, we correlate a standard measure of vertical integration, the industry-level share of intrafirm imports in total U.S. imports, with a new measure of potential cost savings of unethical production for U.S. firms: the share of expenditure on hazardous waste removal in total cost at the industry level. We

[1] Examples include Nike for sweatshops in Indonesia (Harrison and Scorse, 2010); Apple and Samsung for abusive work conditions and environmental pollution in their supplier factories in China (China Labor Watch, 2018, Bloomberg, 2018, and China Labor Watch, 2012); McDonalds, Pepsico, Nestlé, Unilever and Procter and Gamble for rainforest destruction by their palm oil suppliers in Indonesia (Rainforest Action Network, 2017, Guardian, 2017); Coca-Cola for child labor at sugar cane suppliers in El Salvador (Human Rights Watch, 2004, Guardian, 2014); Abercrombie&Fitch, Benetton, C&A, Columbia, Decathlon, Old Navy, Banana Republic, H&M, Levi's, Marks&Spencer, Hilfiger and Calvin Klein for abusive work conditions at Indian supplier factories (Clean Clothes Campaign, 2017) and toy producers Hasbro and Mattel for labor abuse by their suppliers in China (China Labor Watch, 2015, Fortune, 2015). Responding to NGO pressure, the top five apparel brands Nike, Zara, H&M, Adidas, and Uniqlo have - among many others - agreed to remove hazardous chemicals from their entire supply chain by 2020 (Greenpeace, 2016) (ranking according to KantarMillwardBrown, 2017). Baron (2012) and Krautheim and Verdier (2016) provide additional examples.

think of this as a measure of the *potential cost savings* a U.S. firm in this industry could realize in the *absence* of environmental regulation.[2] As potential cost savings can translate into real cost savings only when production takes place in a low-regulation environment, we also allow the effect of this variable to differ systematically across source countries. We therefore include an interaction between our new measure of the potential cost savings of unethical production and the Environmental Policy Stringency Index constructed by the OECD. At the same time, we control for the standard determinants of the international organization of production identified in the previous literature.

Our results are in line with the pattern suggested by the examples above. We find systematic evidence that *a high cost advantage of 'unethical' production in an industry favors international outsourcing* (as opposed to vertical integration) and most strongly so for sourcing from low-regulation countries. In other words: Over and above the well-known determinants of the international integration vs. outsourcing decision, firms in sectors with high potential for 'unethical' cost savings tend to keep suppliers at arm's length rather than to integrate them into the firm. While we identify this *novel* determinant of the international outsourcing decision, we also find that its introduction does not alter the effects of well-known determinants of integration highlighted in other contributions like capital intensity, R&D intensity and skill intensity.

Building on these empirical results, we analyze a model that can rationalize our findings and allows for further theoretical investigation into the underlying mechanisms. The model captures the relevance of both the potential cost savings of unethical production as well as the central determinant of the international organization of production identified in the Antràs literature: the headquarter intensity of production.

In Section 3, we present our baseline model. This model is designed to focus on the effect of (un)ethical production and consumer boycotts on the international organization of production. It delivers our main theoretical results, which directly speak to our empirical findings. In Section 4, we extend our model to provide a foundation for several ad hoc assumptions of the baseline model. There, we introduce elements of the literature on *Private Politics* (Baron, 2003) to show that the results from the analysis of the baseline model are robust to explicitly modeling the NGO-firm interaction leading to boycotts.

We use our baseline model to address the following questions: Why do firms with high potential cost savings of 'unethical' production tend to choose arm's length relationships when sourcing from low-regulation countries? How does the introduction of an 'unethical' technology and consumer boycotts interact with the well-established mechanisms at the heart of the literature on the international organization of production?

In our model, a headquarter in the Global North sources an intermediate input from a supplier located in the Global South. The headquarter can choose between an arm's length relationship (international outsourcing) and vertical integration of the supplier into the firm. The Global South is characterized by a level of regulation and enforcement capacity that allows suppliers to choose between two types of technology: a high-cost and a low-cost technology. The latter delivers identical product quality, but generates an externality on a third party (e.g. local pollution, unsustainable extraction of renewable resources, forced overtime or poor labor, safety and health standards in production).

[2] We focus our attention on a measure of *environmental* cost savings potential as data for a corresponding measure on the labor side are not readily available.

The *first premise* of our analysis is that these externalities raise ethical concerns on the side of consumers in the Global North potentially resulting in a consumer boycott of the final product.[3] We refer to the high-cost and the low-cost technology as 'ethical' and 'unethical', respectively.[4] We assume that the implementation of the unethical technology does not affect the physical properties of the final product. The type of technology used (and therefore the externality it possibly imposes) constitutes a credence attribute: a property consumers do care about, but which cannot be inferred from the final product even after consumption (Feddersen and Gilligan, 2001, and Baron, 2011).

The *second premise* of our analysis is that production along international value chains is characterized by incomplete contracts both in arm's-length relationships *and* in vertically integrated settings. This is the central assumption of the property-rights theory of international production introduced by Antràs (2003). It appears particularly apt in a North-South context where issues concerning dispute settlement, place of jurisdiction as well as questions of enforcement across borders arise. In our context, this contractual incompleteness naturally extends to the implementation of technology: no contract effectively binds the supplier - whether at arm's length or integrated - to implement the ethical or unethical technology type. The massive difficulties of internationally active firms trying to implement codes of conduct for their suppliers backs this assumption.[5]

Like in Antràs (2003), we find that a *high supplier intensity favors international outsourcing* over vertical integration in our model: when the supplier contributes a large part of the production process, also its underinvestment becomes more damaging. The headquarter can then alleviate this underinvestment by keeping the supplier at arm's length.

Moreover, we find that in our model suppliers in sectors with a *high supplier intensity* tend to implement the *unethical* technology. As the technology is not contractible, the supplier implements the type of technology that maximizes its individual expected profits. A high supplier intensity increases the supplier's investment, scaling up the total cost savings of unethical production. This generates a direct link between a high supplier intensity and the implementation of the unethical technology.

Our *first central result* is the identification of a new driver of the integration vs. outsourcing decision that directly interacts with the mechanics of the property-rights theory of the firm. In our model, unethical production increases the gap between the supplier's optimal and *actual* investment. This aggravated underinvestment in the case of unethical production magnifies the incentive of the headquarter to choose outsourcing. We label this the *unethical outsourcing incentive*. It generates a range of supplier intensities for which the headquarter chooses outsourcing if and only if it anticipates unethical production and chooses integration otherwise. In other words, in this range the headquarter keeps the supplier at arm's length *just because* it is unethical.

Our *second central result* is that unethical production by the supplier biases the equilibrium organizational decision of the headquarter towards outsourcing. While this novel link between unethical production

[3]There is ample empirical evidence both from surveys (O'Rourke, 2005, and Loureiro and Lotade, 2005) and from field experiments with real purchasing decisions (e.g. Hiscox and Smyth, 2011, and Hainmueller et al., 2015) that consumers do care about such externalities and have a higher willingness to pay for 'ethical' products. Moreover, Basu and Tzannatos (2003) and Cone (2013) provide evidence that this awareness has increased over the last decades.

[4]We do not take any normative stand on what 'ethical' or 'unethical' practices are. In our analysis, the 'unethical' technology simply features lower costs and imposes an (additional/stronger) externality on a third party.

[5]Nike is a well documented case in point (e.g. Locke et al., 2007b). Other research documenting difficulties of implementing codes of conduct with independent suppliers includes Egels-Zandén (2007), Ruwanpura and Wrigley (2011), and Bird et al. (2019).

and outsourcing is introduced by the unethical outsourcing incentive, the model is rich enough to also generate unethical integration as well as ethical outsourcing as equilibrium outcomes.

Our *third central result* links the model to our empirical findings. We show that the bias towards outsourcing is increasing in the cost advantage of unethical production. That is, everything else equal, firms in sectors with a stronger cost advantage of unethical production are more prone to choosing outsourcing. This implies that our model can indeed rationalize the findings of our explorative empirical analysis: controlling for other sector characteristics - especially the headquarter intensity of production - *firms in sectors with a high potential for unethical cost savings tend to keep suppliers at arm's length rather than to integrate them into the firm.*

Our model also implies an interesting tension between aspirations and reality when it comes to the headquarter's actions. In the public debate, firms are frequently accused of lying when claiming to be in favor of ethical production while actually expanding unethical production. We show that this can be an equilibrium outcome due to the combination of non-contractible technology and the unethical outsourcing incentive. This is the case when the headquarter would *prefer* the supplier to implement the ethical technology (which it cannot impose due to incomplete contracts) but *anticipates* unethical production. The headquarter then has an incentive to maximize cost savings by scaling up unethical production. It can do so by keeping the supplier at arm's length.[6]

In Section 4, we present an extension of the baseline model, in which we explicitly consider the origins of consumer boycotts. This extension provides economic foundations for several ad hoc assumptions in the baseline model and underscores the robustness of our main results. The central premise of this extension is that advocacy NGOs can partially fill the regulatory void by means of private politics. In line with the literature on Private Politics, discussed in detail below, NGOs can investigate the type of technology used and organize campaigns and consumer boycotts to sanction 'unethical' production.[7] They can thereby push firms towards self-regulation i.e. towards implementing the 'ethical' technology although the 'unethical' technology is available.[8] We introduce an NGO screening firms for signs of unethical behavior and organizing consumer boycotts in response. The NGO starts a boycott whenever it is certain that a firm has implemented the unethical technology. This can either be the case when its beliefs (formed according to Bayes' Law) lead the NGO to this conclusion or when a costly investigation of the firm provides evidence of misconduct.[9]

[6]The model also features cases in which the interests of the headquarter and the supplier are aligned and both prefer the ethical or, notably, both prefer the unethical technology.

[7]One of the early and well-documented examples is the case of Nike. In the 1990s, Nike was hit by large-scale protests against work conditions in supplier factories in Indonesia. The annual report for the year 1998, at the height of the protests, states: "[Compared to the prior year,] U.S. footwear, representing the Company's largest market segment, decreased over $255 million in sales [...]. The reduction in sales was primarily attributable to the glut of inventory at retail which reduced customer order volumes and increased order cancellation rates." This was accompanied by a 49.8% decline in profits and the stock price on May 31st, 1998, was 20% below its mark one year earlier (see Nike, 1998). While Nike is an instructive case, several studies provide more systematic evidence that activist campaigns against firms negatively impact their stock market valuation (e.g. King and Soule, 2007, and Flammer, 2013).

[8]In March 1998, Nike CEO Phil Knight made the following statement: "The Nike product has become synonymous with slave wages, forced overtime, and arbitrary abuse. I truly believe the American consumer doesn't want to buy products made under abusive conditions" (New York Times, 1998). Following this statement Nike gave in to NGO demands, see e.g. Harrison and Scorse (2010), and implemented extensive codes of conduct for its suppliers which are analyzed in detail in Locke et al. (2007b).

[9]A key feature of our modeling is that consumers *trust* NGOs, and indeed, surveys repeatedly and consistently find that NGOs are the most trusted institutions before government, business and media (see, for example, the Edelman Trust Barometer for 2019, at `/https://www.edelman.com/sites/g/files/aatuss191/files/2019-03/`

While in the baseline model we impose that failure of an unethical firm to pool with ethical firms leads to a boycott with certainty, we derive this as an equilibrium outcome in the extended model. Moreover, we show that all results from the baseline model hold qualitatively in the extended model, especially those relating to the empirical analysis. By explicitly modeling the NGO-firm interaction potentially leading to consumer boycotts, the extension provides an important private politics underpinning of the baseline model.

Our paper contributes to the large literature on the international organization of production pioneered by Antràs (2003) and surveyed in Antràs and Yeaple (2014). Several determinants of the choice between vertical integration and international outsourcing have been identified and analyzed both theoretically and empirically. This includes capital/headquarter intensity in Antràs (2003), firm heterogeneity in Antràs and Helpman (2004), contractibility of inputs in Antràs and Helpman (2008), task routineness in Costinot et al. (2011), and downstreamness of the supplier in the value chain in Antràs and Chor (2013). In this paper, we suggest the regulatory void, unethical production and consumer boycotts as an additional driver of the international organization of production.[10] The central premise of the property-rights theory of the firm is that the headquarter can never control the investment of the supplier. Be it under integration or outsourcing: contracts always remain incomplete. We extend this to the choice of the production technology. This allows us to stay very close to the Antràs (2003) setting and within the logic of the property-rights theory of the firm. To our knowledge, this is the first paper that introduces the possibility of unethical production into this literature.

Fu et al. (2018) propose an Industrial Organization model of the make-or-buy decision featuring vicarious liability. That is, in case of an infringement, the headquarter faces full liability under integration, but limited liability for an infringement by an independent party. They focus their analysis on the optimal quantities ordered by the headquarter when larger quantities incentivize infringement by the supplier. They do not include the concept of vicarious liability into a model of the international organization of production. Their model is therefore silent on the role of the key determinant of the international organization of production in both the theoretical and the empirical property-rights literature: the headquarter intensity of production. We, in contrast, derive our results from a property-rights setting. Our unethical outsourcing incentive is therefore fully consistent with the logic of the Antràs (2003) framework. This allows us to directly build on, and extend, the empirical branch of this literature to analyze the role of our new variable of interest: the cost advantage of unethical production. Moreover, we can go beyond the mere analysis of our new mechanism to an assessment of how the established theoretical and empirical insights from the Antràs literature are affected when our novel elements and trade-offs are introduced.[11]

`2019_Edelman_Trust_Barometer_Global_Report.pdf`, accessed March 7, 2019. In our model, consumers' trust towards NGOs implies that they are willing to base boycott decisions on recommendations by the NGO. This implies that our modeling of the NGO activity is *fully consistent with incomplete contracts* concerning the technology used. We discuss this point in detail in Section 4.2.

[10] Differences in regulation and institutions are not alien to this literature. For example, Antràs and Helpman (2004), Antràs and Helpman (2008) and Acemoglu et al. (2007) analyze differences in contracting institutions. In contrast to those papers, we do not consider institutions like rule of law and the protection of property rights, but instead consider differences in environmental regulation and labor rights and their enforcement.

[11] We view it as an advantage of our model that it generates the unethical outsourcing incentive from within the mechanics of the property-rights theory of the firm even without imposing vicarious liability. While introducing this into a model of the international organization of production could be an interesting enterprise, it appears questionable to us whether the notion of vicarious liability is consistent with the property-rights theory of the firm, where the headquarter never has control over the supplier's actions.

By explicitly modeling the firm-NGO interaction and the possibility of consumer boycotts in the extended model in Section 4, our paper contributes to bridging the gap between International Economics and the literature on Private Politics started by Baron (2001, 2003). The latter focuses on activists attempting to affect firm behavior not through lobbying for regulation (public politics) but through campaigns and boycotts of firms (private politics). This literature takes an Industrial Organization perspective and analyzes the interaction of activists, firms and possibly a regulator under different market structures and allowing for strategic interactions between all parties.[12]

Most related to our work, a small number of papers have started to introduce elements of Private Politics into International Economics. Aldashev and Verdier (2009) analyze the international competition for funds among development-oriented NGOs. Aldashev et al. (2015) consider the impact of NGO campaigns on industry structure in a setting with endogenous mark-ups and monopolistic competition. Krautheim and Verdier (2016) analyze the endogenous emergence of a consumer-financed NGO in response to the offshoring decision of a firm. Kitzmuller (2012) takes the model of Besley and Ghatak (2007), who explicitly model an NGO as a potential provider of a public good, to the international level.

Most empirical evidence on international NGO campaigns rests on case studies and anecdotal evidence. A notable exception is Hatte and Koenig (2018). Their unique data set contains campaigns of 3359 activists, campaigning against a total of 7168 firms headquartered in 137 countries (for 2010–2015). In addition to underlining the quantitative relevance and the international dimension of the phenomenon, they show that - similar to international trade flows - international NGO campaigns correlate with standard gravity variables. Couttenier and Hatte (2016) shed light on how NGOs adjust their campaigns to media (in)attention, in their case due to large-scale sports events.

Issues related to private regulation, social activism and NGO-firm interactions in global value chains have received much more attention in Political Sciences and Management Studies. This literature finds that social activism is instrumental in the establishment of codes of conduct in multinational supply chains and analyzes further determinants of their success in case studies and, more recently, in large firm-level datasets.[13]

Our work has some relation to several strands of the International Trade literature. First and foremost, Copeland and Taylor (1994) formalize the idea that differences in environmental regulation affect the international location of production. This triggered a large literature on trade (FDI) and the environment which is surveyed in Copeland and Taylor (2004).[14] Our approach has in common with this literature that we view regulatory differences as a driving force of the internationalization of production. This literature, however, does not analyze the international outsourcing vs. integration decision and, importantly, ignores the feedback effects the implementation of unethical technologies can have on demand when consumer boycotts are possible. Moreover, the works of Grossman and Helpman (1994) and Eckel and Egger (2009) have some elements that resonate with our analysis, but address different questions analyzing different

[12]Some of the main contributions include Innes (2006), Baron and Diermeier (2007), Baron (2010), Lyon and Salant (2013) as well as Baron (2016), and Egorov and Harstad (2017). Closely related to the Private Politics literature, but with a different focus, are works on the private provision of public goods and corporate social responsibility (CSR) surveyed by Kitzmuller and Shimshack (2012).

[13]See e.g. Locke et al. (2007a), Distelhorst et al. (2017), Ouellet et al. (2015), Distelhorst and Locke (2018) and references therein.

[14]See Aichele and Felbermayr (2015) and references therein for more recent contributions.

types of actors and strategies.[15]

The remainder of this paper is structured as follows. We present our explorative empirical analysis in Section 2. Section 3 presents our baseline model of unethical production and consumer boycotts and analyzes the optimal international organization of production. In Section 4, we analyze an extension of the model featuring advocacy NGOs and asymmetric information, thereby explicitly modeling the link between firm choices and consumer boycotts. Section 5 concludes.

2 Explorative Empirics

As outlined above, empirical analyses in the property-rights literature on the international organization of production have identified several key determinants of the international integration vs. outsourcing decision. Inspired by the long list of examples in which firms are criticized for actions of their independent suppliers we seek to explore the question whether the (incentive for) unethical production in an industry also contributes to explaining the prevalence of outsourcing in a sector over and above the previously identified determinants. We do so by providing systematic evidence on the relationship between the cost incentives for unethical production and the international organization of production, drawing on data sources and methodology from the established empirical literature. Specifically, our analysis closely follows Nunn and Trefler (2013), Antràs and Chor (2013), Antràs and Yeaple (2014) and Antràs (2016).[16]

2.1 Intrafirm Import Share

We employ the standard measure of vertical integration at the industry level: the share of intrafirm imports in total U.S. imports. Data on intrafirm trade at a detailed country-industry level come from the Related Party Trade Database administered by the U.S. Census Bureau.[17] We use information on U.S. imports in manufacturing from all over the world at the NAICS 6-digit level for the years 2007 to 2014 and convert the data to IO2007 industries from the BEA's input-output tables.

Crucially, the trade flows are distinguished by the relationship between the entities who trade them. A trade flow is marked as taking place between two related parties when the importer holds at least a 6% equity stake in the exporter and as unrelated trade otherwise. We construct our dependent variable, the intrafirm import share, as the value of related party imports over the sum of the value of related and unrelated party imports for each IO2007 industry-country-year.[18] In our main specifications, our regression sample covers around 215 industries.

[15]The 'protection for sale' literature based on Grossman and Helpman (1994) considers the influence of a special interest group on trade policy outcomes. The focus is therefore on *public politics* rather than on *private politics* and on *trade policy* rather than on the *international organization of production*. Eckel and Egger (2009) study the role of trade unions for international investment and production decisions of firms. There are several important differences between advocacy NGOs and trade unions. The former affect firms through demand, tend to be indifferent to survival of the firm and address externalities that usually concern third parties. The latter in turn affect firms on the cost side, vitally depend on firm survival and maximize the utility of their (nationally segmented) members.

[16]We follow the contributions based on industry-level data. There is another strand of this literature that uses firm-level data, e.g. Corcos et al. (2013), Defever and Toubal (2013) and Kohler and Smolka (2014).

[17]The data are available online from `/https://relatedparty.ftd.census.gov/`.

[18]A third category, unreported trade, captures import flows that are not marked as either type of trade. The share of unreported trade flows in total imports is usually negligible. Antràs and Chor (2013) provide a more detailed discussion of the distribution of unreported trade across industries and source countries.

2.2 Measuring the Environmental Unethical Cost Advantage

To the best of our knowledge no direct measures of unethical production exist. We therefore turn to a measure of the cost incentive to implement unethical production, which we label the *unethical cost advantage*. The unethical cost advantage has an industry and a source-country dimension. The industry dimension captures the potential cost savings from lower (or absent) regulation and the country dimension accounts for the actual level of regulation in the respective source country. For reasons of data availability, we focus our analysis on the *environmental* cost incentives of unethical production.

2.2.1 The Industry Dimension

We suggest a measure of potential industry-level cost savings of U.S. firms at the expense of the environment, for which we draw on data from the Annual Survey of Manufactures (ASM) provided by the U.S. Census Bureau. Starting in 2007, the survey records the industry-level expenditure on water, sewer, and refuse removal, as well as other non-electric utility payments including the cost of hazardous waste removal for each year. The measure thus varies over time and across industries. We use this expenditure category as a proxy for the money amount a U.S. firm in that industry would *save* if production took place in an unregulated environment.

One advantage of this measure is that, according to the survey manual, it excludes payments for machinery, equipment, and electric utility.[19] This makes us confident that we capture only those costs that are directly related to the removal of hazardous materials and other waste and that more capital-intensive industries are not mechanically more intensive in waste removal costs.

We construct our variable of an industry's environmental cost savings potential (ECSP) as the log of an industry's expenditure on waste removal *relative to its payroll, total cost, or total sales*, respectively. We will explain in Section 2.3 why the normalization by total costs is our preferred specification.

Table 1 documents the industries with the lowest (left panel) and the highest (right panel) ECSP based on the total cost definition, averaged over the sample period for each industry. In the empirical analysis described below, we will make use of both the industry and the time variation in our variable. Our measure generates a ranking of industries that is arguably in line with common preconceptions about environmentally 'dirty' industries, such as the chemical or textile industries. On the other end of the spectrum our measure puts industries that are mainly involved in assembling parts and thus do not produce a lot of (hazardous) waste.

2.2.2 The Country Dimension

The extent to which the *potential* cost savings translate into *actual* savings depends crucially on the strictness of regulation in the source country. Only if regulation there is more lenient than in the U.S., can (some of) the potential cost savings be realized.

To measure the country dimension of the unethical environmental cost advantage, we employ the Environmental Policy Stringency Index (EPSI) computed by the OECD for 26 member countries (excluding the U.S.) and the six non-member countries Brazil, China, India, Indonesia, Russia, and South Africa for

[19] The survey manual contains detailed instructions about the forms to be filled out by sampled establishments. The manual for the survey year 2015 is available from `/https://bhs.econ.census.gov/bhs/cosasm/ASMInstructions.pdf`. The instruction pertaining to our variable can be found on p.17 of that manual.

Table 1: ECSP Across Industries - Bottom and Top 20

Industry	IO2007 code	ECSP	Industry	IO2007 code	ECSP
334517	Irradiation apparatus manufacturing	0.00055	⋮	⋮	⋮
333112	Lawn and garden equipment manufacturing	0.00060	327100	Clay product and refractory manufacturing	0.00674
336112	Light truck and utility vehicle manufacturing	0.00070	311930	Flavoring syrup and concentrate manufacturing	0.00677
336411	Aircraft manufacturing	0.00076	325411	Medicinal and botanical manufacturing	0.00677
333313	Office machinery manufacturing	0.00081	327910	Abrasive product manufacturing	0.00692
336120	Heavy duty truck manufacturing	0.00083	325320	Pesticide and other agricultural chemical manufacturing	0.00735
336360	Motor vehicle seating and interior trim manufacturing	0.00084	327200	Glass and glass product manufacturing	0.00750
336213	Motor home manufacturing	0.00086	311990	All other food manufacturing	0.00757
311119	Other animal food manufacturing	0.00090	311520	Ice cream and frozen dessert manufacturing	0.00785
334210	Telephone apparatus manufacturing	0.00093	311615	Poultry processing	0.00793
334111	Electronic computer manufacturing	0.00098	332800	Coating, engraving, heat treating and allied activities	0.00815
336999	All other transportation equipment manufacturing	0.00106	313300	Textile and fabric finishing and fabric coating mills	0.00892
334418	Printed circuit assembly (electronic assembly) manufacturing	0.00108	325180	Other basic inorganic chemical manufacturing	0.00944
334220	Broadcast and wireless communications equipment	0.00108	312120	Breweries	0.00954
336212	Truck trailer manufacturing	0.00110	325130	Synthetic dye and pigment manufacturing	0.00957
335314	Relay and industrial control manufacturing	0.00111	322130	Paperboard mills	0.00984
333120	Construction machinery manufacturing	0.00117	327992	Ground or treated mineral and earth manufacturing	0.01041
336211	Motor vehicle body manufacturing	0.00118	327993	Mineral wool manufacturing	0.01334
339910	Jewelry and silverware manufacturing	0.00120	311221	Wet corn milling	0.01339
333993	Packaging machinery manufacturing	0.00121	325120	Industrial gas manufacturing	0.02160
⋮	⋮	⋮	325190	Other basic organic chemical manufacturing	0.02445

Note: The ranking is based on industries for which intrafirm trade data are available. They are ranked by their ECSP measured as expenditure on hazardous waste removal over total cost. Each industry value is an average over 2007-2014. The left panel shows the bottom 20 industries, the right panel the top 20 industries in terms of ECSP.

the years 2007 to 2012. The index combines information on 14 environmental policy instruments that are mainly related to air and climate pollution and is suitable for comparisons across countries.[20] We consider the EPSI a proxy for the general regulatory environment. Implicitly, we thus assume that countries with a low EPSI value are also lenient in their regulation on the removal of water, sewer, refuse, and hazardous waste.

2.3 Control Variables and Intensities

In addition, we use various control variables that have been identified in the reference literature as determinants of intrafirm trade or have been used for robustness checks therein. To control for headquarter intensity (the counterpart of supplier intensity) we include the logs of capital intensity, R&D intensity and high-skill intensity. We also control for material intensity (normalized expenditure on materials), within-industry size dispersion and the elasticity of substitution. We follow the literature and disaggregate capital into its components, which arguably differ in relationship-specificity, to obtain a cleaner proxy for headquarter intensity.[21]

In our preferred specifications we construct all intensities as the log of the respective expenditure relative to total industry cost (for R&D we follow the literature and normalize by sales). We construct a proxy for total cost, for which we sum payroll, cost of materials, total capital expenditure, total rental payments and an aggregate term for all other expenditures from the ASM. We assess the robustness of

[20] According to the OECD's definition, a policy is more stringent if it puts a higher explicit or implicit price on pollution or environmentally harmful behavior. An index value of 0 is the lowest stringency possible, while an index value of 6 denotes the highest stringency. The maximum value the index attains in our sample is 4.41 for Denmark in 2009. The lowest value is .375 for Brazil in 2011.

[21] All data come from standard sources. Details on these sources and the construction of the variables can be found in the Online Appendix. In using U.S. data to control for industry characteristics we maintain the assumption that relative factor intensities are correlated across countries, as is standard in this literature. See e.g. Nunn and Trefler (2013) or Costinot et al. (2011).

our results by normalizing with total sales as well. For a direct comparison to the literature we also report results for a normalization with payroll. We prefer the total cost definition because we believe it captures most directly the relative importance of a cost category, and in particular that of the hazardous waste removal costs, in the overall production process.[22]

2.4 Empirical Specification

We estimate variants of the following regression equation.

$$intrafirm_{ijt} = \eta_0 + \eta_1\ ECSP_{jt} + \eta_2\,(ECSP_{jt} \times EPSI_{it}) + \rho X_{jt} + \zeta_{it} + \epsilon_{ijt}. \tag{1}$$

$intrafirm_{ijt}$ is the share of related-party imports in total imports by the U.S. from country i in industry j in year t. $ECSP_{jt}$ is our proxy for the part of the *unethical environmental cost advantage* varying across industries j and over time t. $EPSI_{it}$ proxies for the part of the *unethical environmental cost advantage* that varies across source countries i and time t. X_{jt} contains the established determinants of intrafirm trade and the other control variables mentioned above. ζ_{it} is a set of country-year fixed effects to control for everything that is specific to a country in a given year. The fixed effects therefore control for the level effect of $EPSI_{it}$. They also control for the endogenous choice of a sourcing location to the extent that this is driven by country- or country-year-specific factors, such as geography, corporate tax rates or cultural linkages. We want to take out this variation to be able to make statements about the tendency to outsource production conditional on the chosen source country. In all our regressions, we cluster standard errors at the IO2007-industry level as this is the level of variation of our main explanatory variables and industry characteristics are highly auto-correlated over time.

Our data on intrafirm imports cover 230 countries and territories. But our measure of the level of regulation, EPSI, is limited to 26 OECD countries (excluding the U.S.) plus the six non-member countries listed above. We therefore run the specification in equation (1) in two versions.

Making use of the full sample, we only include $ECSP_{jt}$ but not the interaction effect in the first specification. In support of the anecdotal evidence presented in the Introduction, we would expect $\eta_1 < 0$: industries with higher potential cost savings should have a lower share of intrafirm trade, indicating a link between (incentives for) unethical production and outsourcing.

We expect this under the premise that most of the 230 countries and territories have more lenient regulation and enforcement (capacity) than the U.S. Within the set of countries for which EPSI data are available the U.S. takes a middle position. Arguably, many, if not most, of the 198 countries and territories for which EPSI is *not* available (especially most of the remaining non-OECD countries) should indeed be expected to have more lenient regulation and enforcement (capacity) than the U.S. The presence of countries with similar or higher levels of regulation should bias the results against our hypothesis, so it is safe to keep them in the sample.

[22] An additional argument for using broader measures of factor intensities is the fact that the share of capital expenditure in total cost and the share of the wage bill (payroll) in total cost are significantly correlated with a positive coefficient of 0.1345 in our data. The correlation coefficient is 0.1687 when using total sales in the denominator. This correlation is puzzling when one has a Cobb-Douglas production function in mind with labor and capital as inputs. In the data, a very large portion of an industry's expenditure is allocated to intermediate inputs. When we correlate the sum of payroll and material input expenditure relative to total cost with the share of capital expenditure in total cost, the correlation coefficient is highly significant at -0.5677. We therefore prefer to compute the intensities using total cost in the denominator.

In the second specification, we add the interaction of the cost savings potential $ECSP_{jt}$ and $EPSI_{it}$, the OECD Environmental Stringency Index. Due to the limited coverage of the $EPSI_{it}$ we have a strongly reduced sample size in this specification. On the other hand, the interaction effect allows for more flexibility to analyze the differential impact of a given level of $ECSP_{jt}$ across varying regulatory environments. The tendency to outsource production in industries with a given $ECSP_{jt}$ should be stronger when the goods are sourced from countries with more lenient environmental policies. In the second specification, we therefore expect $\eta_2 > 0$ and continue to expect $\eta_1 < 0$.

2.5 Empirical Results

In this subsection we present our estimation results. In our preferred specification, we normalize the explanatory variable with total cost. We then show that the results we find also hold qualitatively when we normalize with total sales and payroll. First, however, we show that the well-established results in this literature also hold in our data.

Columns 1, 4, and 7 of Table 2 show the typical regression specification in the reference literature. Using the payroll normalization, column 7 shows that we can qualitatively replicate the main results of the literature. Columns 1 and 4 show that the change in normalization does not affect these findings.

In column 2 of Table 2 we add our measure of the environmental cost savings potential (ECSP) to our preferred specification with the total cost normalization. We indeed find that a larger ECSP is associated with less intrafirm trade on average and is significant at the 10% level. The other coefficients do not change much compared to column 1 and continue to have the right signs. Industries with a higher ECSP outsource more. We also calculate the standardized beta coefficient associated with the estimate and find that the intrafirm trade share decreases by 4.5% of a standard deviation when the log of the ECSP increases by one standard deviation. The corresponding beta coefficient for the log of other machinery intensity, the standard proxy for relationship-specific headquarter intensity, implies an increase in the intrafirm trade share by 6.3% of a standard deviation. We take this as evidence of the economic significance of our new variable.

In column 3 we add the interaction term of the ECSP with the index of environmental policy stringency (EPSI). Consistent with our expectations, we find the interaction effect to be positive and significant at the 5%-level. The level effect of the ECSP almost doubles in absolute magnitude and is negative and significant at the 1%- level. The interaction effect uncovers a strong cross-country pattern of heterogeneity in the effect of the ECSP.[23]

To analyze the cross-country dimension further, we report marginal effects of the ECSP at various percentiles of the distribution of the EPSI. In Table 3, column 1 shows the marginal effects and the corresponding p-values for the total cost specification from column 3 of Table 2. There is sizable variation in the marginal effects. The coefficients are significant at the 1%-level up to and including the first decile. The coefficients turn insignificant by conventional levels at the fourth decile. The magnitude of the

[23]Due to to the limited coverage of the EPSI our sample drops to roughly one fourth of its former size as we have to exclude the many non-OECD countries (except the six emerging economies mentioned above) for which we do not have data. In the Online Appendix we provide additional tables which show that the level effect of the ECSP is also negative when we remove the interaction effect and hold the (small) sample size constant. In many cases the level effect is not significant when the sample size is reduced, indicating that, in line with our expectations, it is indeed countries outside the realm of developed OECD countries driving our results in Columns 2, 5, and 8.

Table 2: The Effect of Unethical Environmental Cost Advantage on Intrafirm Trade

Dependent Variable: Intrafirm Import Share									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intensity Definition:	Total Cost	Total Cost	Total Cost	Total Sales	Total Sales	Total Sales	Payroll	Payroll	Payroll
log ECSP		-0.0228*	-0.0401***		-0.0204*	-0.0388***		-0.0241**	-0.0270*
		(0.0123)	(0.0143)		(0.0121)	(0.0143)		(0.0116)	(0.0140)
log ECSP			0.00893**			0.00917**			0.00174
X EPSI			(0.00428)			(0.00434)			(0.00410)
log other machinery	0.0286**	0.0385***	0.0560***	0.0299***	0.0387***	0.0503***	0.0263***	0.0389***	0.0488***
intensity	(0.0114)	(0.0111)	(0.0140)	(0.0101)	(0.0103)	(0.0145)	(0.00972)	(0.0101)	(0.0137)
log skill intensity	0.0392*	0.0368*	0.0584***	0.0344***	0.0341***	0.0375**	0.0481*	0.0508*	0.0491
	(0.0221)	(0.0212)	(0.0211)	(0.0121)	(0.0118)	(0.0152)	(0.0282)	(0.0276)	(0.0375)
log R&D intensity	0.0221***	0.0207***	0.0267***	0.0223***	0.0210***	0.0278***	0.0214***	0.0197***	0.0269***
	(0.00392)	(0.00405)	(0.00481)	(0.00383)	(0.00396)	(0.00494)	(0.00450)	(0.00452)	(0.00543)
log materials	0.0744	0.0588	0.130**	0.0548**	0.0481*	0.0455	-0.00465	0.00383	-0.00968
intensity	(0.0634)	(0.0624)	(0.0590)	(0.0225)	(0.0246)	(0.0330)	(0.0116)	(0.0115)	(0.0116)
dispersion	0.0840***	0.0789***	0.0862***	0.0836***	0.0791***	0.0871***	0.0821***	0.0766***	0.0850***
	(0.0138)	(0.0133)	(0.0139)	(0.0139)	(0.0135)	(0.0146)	(0.0137)	(0.0131)	(0.0147)
log building	-0.0101	-0.00815	-0.0117	-0.00815	-0.00678	-0.0103	-0.0131**	-0.0105*	-0.0146*
intensity	(0.00635)	(0.00634)	(0.00771)	(0.00636)	(0.00629)	(0.00740)	(0.00602)	(0.00617)	(0.00774)
log auto intensity	-0.0117**	-0.0120***	-0.0182***	-0.0134***	-0.0133***	-0.0211***	-0.0108**	-0.0109**	-0.0182***
	(0.00458)	(0.00438)	(0.00597)	(0.00459)	(0.00444)	(0.00615)	(0.00472)	(0.00444)	(0.00629)
log computer	-0.00909	-0.0121	0.000720	-0.00693	-0.0101	0.00337	-0.0116	-0.0147*	-0.00233
intensity	(0.00767)	(0.00781)	(0.0106)	(0.00714)	(0.00754)	(0.0108)	(0.00770)	(0.00766)	(0.0109)
sigma	-0.000373	-0.000554	-4.90e-05	-0.000195	-0.000387	-1.10e-05	-0.000366	-0.000544	-0.000110
	(0.000498)	(0.000497)	(0.000619)	(0.000441)	(0.000449)	(0.000594)	(0.000501)	(0.000498)	(0.000643)
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
IO2007 clusters	217	215	212	217	215	212	217	215	212
Observations	130,947	130,364	35,434	130,947	130,364	35,434	130,947	130,364	35,434
Adj. R-squared	0.168	0.170	0.164	0.169	0.171	0.164	0.168	0.170	0.162

Note: Estimation by OLS with standard errors clustered at the industry level reported in parentheses. ***, **, and * denote significance the 1%, 5%, and 10% level, respectively. log ECSP is the log of expenditure on waste and hazardous materials removal over payroll, total cost or total sales. sigma is the estimate of the import demand elasticity from Broda and Weinstein (2006).

marginal effect continues to fall until it reaches a value of almost zero at the ninth decile of our sample. Table 3 clearly shows that the effect of the ECSP on intrafirm trade is driven by the countries with the lowest environmental regulation. This underscores the importance of regulatory differences across countries in shaping the effect of potential unethical cost savings on the international organization of production.[24]

Columns 4 to 6 of Table 2 show that our result is not driven by the normalization variable we use. The coefficient of the ECSP is weakly significant and negative in column 5, and becomes larger in absolute terms and highly significant when we add the interaction term, which is positive and significant at the 5%-level here as well. Turning to the marginal effects for the total sales normalization in column 2 of Table 3, it is evident that the pattern is qualitatively and quantitatively almost identical to the one from the total cost specification.

[24]We have also run regressions leaving out countries with a stricter EPSI value than the U.S. as well as dropping all OECD members from the sample (not reported). As expected, this changes significance levels and coefficient values of the level effect of ECSP mildly in favor of our hypothesis.

Table 3: Marginal Effects of the ECSP

Intensity Definition	(1) Total Cost		(2) Total Sales		(3) Payroll	
EPSI percentile	marginal Effect	p-value	marginal Effect	p-value	marginal Effect	p-value
5	-0.036	0.007	-0.034	0.009	-0.026	0.043
10	-0.035	0.008	-0.033	0.010	-0.026	0.041
20	-0.027	0.024	-0.025	0.032	-0.024	0.035
30	-0.022	0.060	-0.020	0.079	-0.024	0.039
40	-0.020	0.102	-0.018	0.134	-0.023	0.045
50	-0.017	0.152	-0.015	0.197	-0.023	0.052
60	-0.016	0.203	-0.014	0.260	-0.022	0.059
70	-0.014	0.275	-0.012	0.346	-0.022	0.070
80	-0.012	0.373	-0.009	0.461	-0.021	0.085
90	-0.010	0.446	-0.008	0.545	-0.021	0.096

Note: Marginal effects of log ECSP at deciles of the EPSI are calculated from the regressions in Table 2, columns 3, 6, and 9, respectively.

Columns 7 to 9 of Table 2 provide evidence using the payroll definition of intensities. When included by itself in column 8, the effect of the ECSP is negative with roughly the same magnitude as the coefficients from columns 2 and 5. It is significant at the 5%-level. When we add the interaction effect in column 9, the pattern holds qualitatively, with a negative level effect and a positive interaction term. However, significance levels are lower than in the other specifications. This result is reflected in column 3 of Table 3. The magnitude of the marginal effect changes only very little over the distribution of the EPSI while significance levels range from 5% below the median and a 10%-level of significance up to the ninth decile.

2.6 Robustness

In the Online Appendix, we show that our results continue to hold in a large number of robustness checks. In particular, our results continue to hold when we add a measure of downstreamness and its interaction with the elasticity of substitution as in Antràs and Chor (2013) and include further industry-level controls used in that paper. Importantly, we show that the interaction effect between ECSP and EPSI is robust to using industry-year fixed effects and adding further country-industry interaction terms controlling for institutional and endowment-based sources of comparative advantage.

2.7 Summary and Next Steps

The results of this explorative empirical analysis show that the cost incentive to produce unethically plays a statistically and economically significant role in the international organization of production. A stronger cost advantage of unethical production favors outsourcing. This effect is stronger for source countries with a weak regulatory environment.

These to our knowledge novel results show that there is a systematic relationship between international outsourcing and the incentives for unethical production. This raises further questions, such as: Why do firms with high potential cost savings of 'unethical' production tend to choose arm's length relationships when sourcing from low-regulation countries? How does the introduction of an 'unethical' technology and consumer boycotts interact with the well-established mechanisms at the heart of the literature on the international organization of production? In the following section, we outline a model of the international

organization of production based on Antràs (2003) that includes (un)ethical production and consumer boycotts. The model rationalizes the empirical findings and allows us to address these additional research questions. In Section 4, we analyze an extension of the model featuring advocacy NGOs and asymmetric information, thereby explicitly modeling the link between firm choices and consumer boycotts.

3 Baseline Model

In this section, we outline and analyze our baseline model.

3.1 Setup

We begin by describing the setup of the model, including preferences, technology, as well as our central new elements: consumer boycotts and unethical production.

3.1.1 Preferences, Consumer Boycotts and Demand

All consumers are located in the Global North. Their preferences are summarized by the following CES aggregate over a large number of symmetric varieties indexed by ω,

$$U = \left(\int_{\omega \in \Omega} y(\omega)^{\alpha} I(\omega) d\omega \right)^{\frac{1}{\alpha}}, \tag{2}$$

with $\alpha \in (0,1)$, Ω being the set of available varieties and $y(\omega)$ representing the quantity consumed of variety ω. These preferences are standard with the exception of the indicator variable $I(\omega)$. It reflects the fact that a firm (and its variety) can be hit by a consumer boycott. In this case the indicator variable takes a value of zero implying that the representative consumer does not derive any utility from its consumption.

Consumers maximize their utility subject to the budget constraint $E \leq \int_{\omega \in \Omega} p(\omega) y(\omega) d\omega$. Therefore, in general, demand for each variety ω is given by

$$y(\omega) = A p(\omega)^{-\frac{1}{1-\alpha}} I(\omega)^{\frac{1}{1-\alpha}}, \tag{3}$$

where $A = E \left(\int_{\omega \in \Omega} p(\omega)^{-\frac{\alpha}{1-\alpha}} I(\omega)^{\frac{1}{1-\alpha}} d\omega \right)^{-1}$.

From equation (3) we can see how demand responds to a boycott. In this case the indicator variable takes the value of zero and there is no demand for the product.[25] The value of the preference shifter depends on the choice of the firm and nature (in the baseline model) or the activity of an advocacy NGO (in the extended model). This stylized assumption allows us to generate the risk of losing final revenues as a consequence of unethical production in a simple way that preserves tractability of the model.[26]

[25] Technically, this modeling is a variation of the standard approach in the literature on quality and international trade with CES preferences, where firms can invest into quality represented by a (usually continuous) variable which takes the place of our indicator variable. See e.g. Hummels and Klenow (2005) and Hallak (2006) for early contributions. Baldwin and Harrigan (2011) provide a review of this literature.

[26] Seeing demand drop to zero in case of a boycott may appear to be a strong assumption, as even very successful campaigns usually do not lead to zero demand. For the mechanisms of the model to work, it is key that a boycott leads to *a* reduction in demand. Modeling this reduction in the starkest possible way, i.e. as a drop to zero, increases tractability of the model. This is, however, without compromising the mechanics of the model. We will see below that for the supplier's choice between the ethical and unethical technology *expected profits* are key. These are determined by *demand* with and without a boycott

3.1.2 Production of the Final Good and the Intermediate Input

We focus on the decisions taken by a single firm that is active in a sector with a continuum of homogeneous firms operating under monopolistic competition.[27] The final good is produced by a headquarter located in the Global North using an intermediate good provided by a supplier located in the Global South. The headquarter can costlessly transform one unit of an intermediate good into final output:

$$y(\omega) = x(\omega). \tag{4}$$

The quantity $y(\omega)$ of the final good is therefore simply given by the quantity $x(\omega)$ of the intermediate good the headquarter has at its disposal. The intermediate good is in turn produced by a supplier combining a headquarter service and a manufacturing input according to the following production function:

$$x(\omega) = \left(\frac{h(\omega)}{\beta}\right)^{\beta} \left(\frac{m(\omega)}{1-\beta}\right)^{1-\beta}, \tag{5}$$

where $\beta \in (0,1)$ is the sector-specific headquarter intensity of production (and $1-\beta$ is the supplier intensity). The headquarter service $h(\omega)$ is provided to the supplier by the headquarter which then combines it with the manufacturing input $m(\omega)$ to produce $x(\omega)$ units of the intermediate good. The intermediate good produced is entirely relationship-specific. Neither can the supplier sell $x(\omega)$ to any third party nor can the headquarter produce any of the final output without the intermediate good that is in the possession of the supplier.

We stress that the manufacturing input $m(\omega)$ stands for a bundle of factors of production used by the supplier. Among these are labor and physical capital, as well as human capital and materials. In addition, and crucially for our model, the supplier also incurs other expenditures, such as provisions for workplace safety and the cost of compliance with local environmental regulation in the process of providing the input $m(\omega)$.[28]

3.1.3 Unethical Production and Consumer Boycotts

The central innovation in this paper is that the supplier does not only choose the investment necessary to produce the manufacturing input, but can also choose between a high- and a low-cost technology. The low-cost technology produces a (higher) negative externality on a third party. We can think of such externalities as taking the form of exploitation of workers with forced overtime, low work safety standards or child labor as well as pollution of the environment, e.g. by dumping dangerous chemicals in

combined with the *probability of a boycott*. This boycott probability can take any value between zero and one. This implies that even with a boycott leading to zero demand, our model can generate *any* value of expected profits that a model with positive demand under a boycott could deliver. In light of this argument the probability of a boycott introduced below may also be interpreted in a broader sense as the propensity of a sector to suffer from a boycott.

[27] This allows us to highlight in the starkest possible way how the introduction of an unethical technology and consumer boycotts affects the fundamental results and mechanics of Antràs (2003). Two comments are in order. First, we explain in Section 3.3.3 why also a model with homogeneous firms is consistent with the empirical approach taken in Section 2 of using the *fraction* of intrafirm trade of a sector as dependent variable. This is *despite* the fact that, as we will see below, with homogeneous firms, all firms in a sector take the same decision regarding the organization of production and the (un)ethical technology choice. Second, the attentive reader of our extended model spelled out in Section 4 might think that the link between investment choices and boycotts crucially depends on the assumption of homogeneous firms. We explain in Footnote 41 why this is not the case.

[28] This interpretation is consistent with recent contributions to this literature, such as Antràs and Yeaple (2014).

rivers, emitting substantial quantities of carbon dioxide or harvesting old growth rainforests. Consumers in North consider an unnecessarily high (but cost-saving) level of this externality as unethical. We define the marginal cost of the supplier's high-cost, ethical technology as c_m^e and the low-cost, unethical technology by $c_m^u = \mu c_m^e$, with $\mu \in (0,1)$.[29]

As consumers cannot infer from the final product whether the unethical technology was used in production, unethical firms can potentially prevent consumers from learning about the type of the firm. While we assume that the technology used cannot be directly observed by consumers, some firm choices are observable, potentially leading consumers to believe that the firm is of the unethical type. In the baseline version of the model we follow Krautheim and Verdier (2016) and impose a simple link between observable choices (investments, quantities and prices) of the firm and the probability of facing a boycott: an unethical firm setting observables at values that are optimal for an ethical firm ('mimicking') has a chance to pass as an ethical firm and faces a boycott with probability $1 - \gamma < 1$. *Any firm deviating from the investments, quantities or prices of ethical firms faces a boycott with probability one.*In fact, we only need to impose this for investments, as conditional on identical investments, the same quantities and prices maximize profits of both firm types.[30] This implies that an unethical firm faces discontinuous demand being positive in expectation if and only if it chooses investments like an ethical firm. This leads unethical firms to *mimic* ethical firms and in equilibrium the levels of investment, quantities and prices do not reveal the type of the firm.

Using this reduced-form approach in the baseline model allows us to focus on the analysis of the international organization of production with unethical technologies, to derive our main results on the integration and technology decision, their interaction as well as empirical implications. However, the reduced-form approach leaves some questions open: What is the mechanism/the agent triggering a consumer boycott? Why is it triggered by a deviation from ethical firm choices? Should ethical firms adjust their investments in order to signal their type? To address these questions, we provide an extended version of the model in Section 4, where consumer boycotts are initiated by an NGO that can investigate firms. In the baseline model it is simply imposed that failure of an unethical firm to pool with ethical firms leads to a boycott with certainty. In the extension, we explicitly model how the NGO's formation of beliefs of the firm type conditional on observables delivers this result as an equilibrium outcome.

3.1.4 Hold-up Problem and the Organization of the Firm

We consider an environment with incomplete contracts. Neither can contracts be written contingent on choices the parties make, nor on outcomes like revenue. The only contractible items are the lump-sum transfer from the supplier to the headquarter (discussed in detail below) and the organization of production. This means that investment quantities are not contractible, but also that our new feature,

[29] We do not take any *normative* stand on what an ethical or unethical technology is. This includes for example the debate on the desirability of a ban of child labor. We simply start our modeling from the observation that consumer boycotts are triggered by the perception of (some) consumers that firms act in an unethical way. Clearly, what is considered 'unethical' may depend on the historical context, income, culture, salience of specific issues in the public debate as well as alternative technologies.

[30] This is because after investments are made, the 'optimization' of a firm with respect to quantities and prices is equivalent to a situation in which all costs are sunk, marginal costs are zero, and the maximum output is fixed and identical for both firm types as investments are the same. Therefore, when both firms have set the same investment levels and there is positive demand for the unethical firm, both firms will set the same quantities and market clearing prices.

the technology choice of the supplier, cannot be contracted upon.[31]

As contracts are incomplete, neither the investments nor the split of the revenues can be fixed ex-ante. The relationship-specificity of investments then implies that after investments are sunk and the intermediate input is produced, the two parties face a hold-up problem. Both parties need the partner in order to generate (full) revenue and therefore engage in a bargaining process over the split of final revenues. Following the literature, we model this ex-post bargaining as generalized Nash bargaining with the headquarter getting a fraction of the final revenues. This fraction is endogenous and depends on the residual rights of control, which are in turn affected by the organization of production chosen by the headquarter.

Before investments take place, the headquarter can choose between integrating the supplier into the firm or leaving it as an independent party. We index the mode of organization by $k \in \{O, V\}$, where O stands for international outsourcing and V for vertical integration. The key difference between the two is that outsourcing leaves the supplier with the residual rights of control over the produced intermediate. In this case the outside options of both parties are zero if bargaining fails: the headquarter has no input to produce the final product and the supplier cannot transform the intermediate into the final product.

Integration in turn shifts the residual rights of control to the headquarter allowing it to recover a fraction $\delta \in (0, 1)$ of the intermediates from the supplier if bargaining fails. The outside option of the headquarter under integration is therefore better than under outsourcing, implying that the bargaining results in a larger share of revenues going to the headquarter, i.e. $\phi_V > \phi_O$, where, as in Antràs (2003), $\phi_V = \phi_O + \delta^\alpha (1 - \phi_O)$.

We assume $\phi_k > \frac{1}{2}$. Antràs (2003) shows that this assumption is sufficient to ensure that the headquarter optimally produces the headquarter service by itself and hands it over to the supplier for production of intermediate $x(\omega)$ while the supplier produces the manufacturing input.

3.1.5 Match Creation and Transfer Payment

We have now described the situation after a headquarter has been matched to a supplier. Ex-ante, the headquarter faces a large number of perfectly competitive suppliers available for a match. Once a match is formed, their relationship is transformed into one of bilateral monopoly (Williamson, 1985) in that investments are relationship-specific and have no outside value. Due to incomplete contracts, the production process involving bargaining over the revenues will leave the supplier with positive profits. The large number of potential suppliers compete for this profitable opportunity by offering a transfer payment to the headquarter in return for forming the match with them. Perfect competition among suppliers implies that the headquarter can set a payment that extracts the full expected surplus from the supplier. Besides the organization of production, the transfer payment is the only variable the headquarter and supplier can contract on. Both are fixed in the moment the match is formed.

[31] Alternatively, one could assume that integration allows the headquarter to *impose* the technology on the supplier. This would, however, mix property-rights theory (for production) and the transaction cost approach à la Grossman and Helpman (2002) or Carluccio and Bas (2015) (for technology). In the latter, all contractual incompleteness is resolved by integration. It appears hard to justify the assumption that under integration the headquarter *can* impose the type of technology but *cannot* impose the level of investment. More interesting might be the analysis of a setting that fully embraces the logic of the transaction cost approach where integration allows the headquarter to impose both the investment and the technology. We leave this alternative model for future research and focus in this paper on the predominant paradigm in the literature: the property-rights theory of the firm.

3.1.6 Time Line

Figure 1 gives an overview of the sequence of events. In t_0, the headquarter chooses the organizational form and the lump-sum transfer. In $t_1(a)$, the supplier chooses between ethical and unethical production. Both parties make their physical investments non-cooperatively in $t_1(b)$. The headquarter hands the headquarter service to the supplier, who in turn produces intermediate inputs in t_2 by combining the headquarter service with its own manufacturing input. In t_3, nature determines whether an unethical firm will be boycotted by consumers. Period t_4 features the ex-post bargaining over the division of the surplus. In t_5, if the parties have agreed on a division, intermediates are converted to final output, sold and revenues distributed to headquarter and supplier if the firm is not boycotted. In case of a boycott, demand is zero and no final goods are produced and sold.

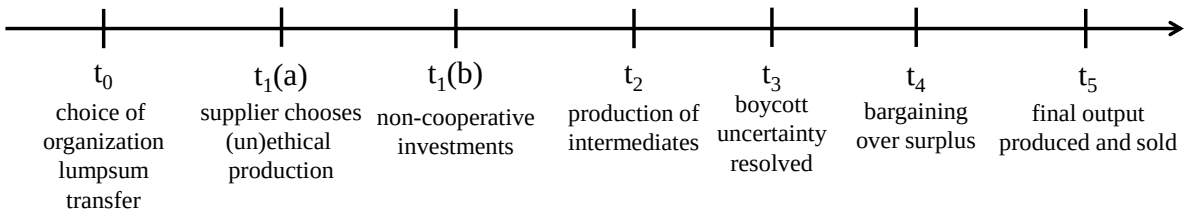

Figure 1: Timing of Events.

3.2 Equilibrium Firm Choices

The focus of our analysis is on two key firm choices: the organization of production and the choice of the (un)ethical technology. We focus on the decisions taken by a single firm. The results, however, extend to the sector as a whole.[32] We solve the model by backward induction.

3.2.1 t_5: Revenues of Ethical and Unethical Firms

We denote revenue from selling variety ω as $R(\omega)_k^l$, where $k \in \{V, O\}$ indicates vertical integration and outsourcing and $l \in \{e, u\}$ indicates ethical and unethical production. An ethical firm always faces full demand as it is never targeted by a consumer boycott. Its revenues are given by $R(\omega)_k^e = p(\omega)_k^e \; y(\omega)_k^e$. $h(\omega)_k^e$ and $m(\omega)_k^e$ represent the investment quantities chosen by headquarter and supplier in the case of ethical production. Given that the quantity $x(\omega)$ of the intermediate good produced by the supplier is determined by investments and given that the headquarter costlessly transforms $x(\omega)$ into $y(\omega)$, total

[32] Recall that we consider a single firm active in a sector with a continuum of homogeneous firms. In taking its decisions, the firm takes aggregate variables like aggregate expenditure and the price index as given. As this holds true for all firms in a sector and as all firms face the same optimization problem, a single firm's choices regarding organization and technology extend to the sector as a whole. Note that this implies that our theoretical results are independent of the industry equilibrium. We argue in Section 3.3.3 that this also holds for the link of the theory to the empirical findings. We do not analyze the industry equilibrium. This would allow us to analyze some additional objects (like the number of firms, aggregate output in a sector or welfare effects). These do not, however, affect the choice of organization and technology in a sector and are therefore orthogonal to the focus of our analysis.

revenues of an ethical firm can be expressed as

$$R(\omega)_k^e = A^{1-\alpha}\left[\left(\frac{h(\omega)_k^e}{\beta}\right)^{\beta}\left(\frac{m(\omega)_k^e}{1-\beta}\right)^{1-\beta}\right]^{\alpha}. \tag{6}$$

An unethical firm only faces positive demand in expectation if $h(\omega)_k^u = h(\omega)_k^e$ and $m(\omega)_k^u = m(\omega)_k^e$. Its revenues under mimicking and if it does not face an exogenous boycott in t_3 are also given by the above expression.

3.2.2 t_4: Bargaining

Headquarter and supplier bargain over the distribution of revenue. The bargaining power - and therefore also the share of revenue - of the headquarter is assumed to be $\phi_O > \frac{1}{2}$ under outsourcing. This reflects the fact that in the arm's length relationship, both parties have an outside option of zero and the payoff allocation is determined only by the exogenous assumptions about the distribution of the gains from trade. In the case of integration, the outside option of the supplier remains at zero because of the relationship-specificity of the produced intermediates. The headquarter, however, has allocated the residual rights of control to itself. It is able to continue producing $\delta y(\omega)$ in case bargaining breaks down. Using equations (4), (5), and (6) this translates into sales of $\delta^{\alpha} R(\omega)_k^l$. The gains from trade are thus reduced to $(1-\delta^{\alpha})\, R(\omega)_k^l$. With integration, the headquarter receives its larger outside option plus its exogenous share from the gains from trade, which is $\phi_V R(\omega)_k^l$, with ϕ_V as defined in Section 3.1.4.

3.2.3 t_3 and t_2: Consumer Boycotts and Production of Intermediates

In period t_3, nature decides whether an unethical firm faces a boycott. We assume that ethical firms never face a boycott, firms that are openly unethical always face a boycott and firms that mimic ethical firms in terms of prices, output, and investment face a boycott with a probability $1-\gamma$. Before the boycott uncertainty is resolved, a mimicking unethical firm therefore has an expected revenue of

$$E[R(\omega)_k^u] = \gamma R(\omega)_k^e. \tag{7}$$

In period t_2, the supplier uses the invested quantities to produce intermediate output $x(\omega)$. As outlined above, provided it mimicked in terms of investments in t_1, there is no reason for an unethical firm to deviate from the optimal quantity of an ethical firm, which is production according to equation (5).

3.2.4 $t_1(b)$: Investments

Two types of decisions are taken sequentially in period t_1. In period $t_1(a)$ the supplier chooses to implement the ethical or unethical technology. In period $t_1(b)$ supplier and headquarter take their investment decisions simultaneously. We first consider the investment choices conditional on the ethical or unethical technology being implemented.

Ethical Investments: When the supplier implements the ethical technology, the setting is isomorphic to Antràs (2003). The two parties simultaneously and non-cooperatively set investments to maximize their respective shares of final revenue. They take into account incomplete contracts and the ensuing ex-post bargaining. The headquarter maximizes

$$\max_{h(\omega)_k^e} \phi_k R(\omega)_k^e - c_h h(\omega)_k^e, \tag{8}$$

whereas the supplier solves

$$\max_{m(\omega)_k^e} (1-\phi_k) R(\omega)_k^e - c_m^e m(\omega)_k^e. \tag{9}$$

Notice the superscript in the marginal cost of the supplier. With ethical production, the supplier rewards its factor of production at the ethical rate c_m^e.

The first order conditions deliver the best response functions that give optimal investment of each party for any positive level of investment of the other party:

$$h(\omega)_k^e = \beta \left(\frac{\phi_k \alpha}{c_h}\right)^{\frac{1}{1-\beta\alpha}} A^{\frac{1-\alpha}{1-\alpha\beta}} \left(\frac{m(\omega)_k^e}{1-\beta}\right)^{\frac{(1-\beta)\alpha}{1-\beta\alpha}}$$

$$m(\omega)_k^e = (1-\beta) \left(\frac{(1-\phi_k)\alpha}{c_m^e}\right)^{\frac{1}{1-(1-\beta)\alpha}} A^{\frac{1-\alpha}{1-(1-\beta)\alpha}} \left(\frac{h(\omega)_k^e}{\beta}\right)^{\frac{\beta\alpha}{1-(1-\beta)\alpha}}.$$

Like in Antràs (2003), the equilibrium of the investment game is at the intersection of the best response functions. The standard argument of Pareto-dominance rules out the other Nash equilibrium at zero-zero. Equilibrium investments are therefore given by

$$h(\omega)_k^e = \beta A \alpha^{\frac{1}{1-\alpha}} \frac{\phi_k}{c_h} \left[\left(\frac{c_h}{\phi_k}\right)^{\beta} \left(\frac{c_m^e}{1-\phi_k}\right)^{1-\beta}\right]^{\frac{-\alpha}{1-\alpha}} \tag{10}$$

$$m(\omega)_k^e = (1-\beta) A \alpha^{\frac{1}{1-\alpha}} \frac{1-\phi_k}{c_m^e} \left[\left(\frac{c_h}{\phi_k}\right)^{\beta} \left(\frac{c_m^e}{1-\phi_k}\right)^{1-\beta}\right]^{\frac{-\alpha}{1-\alpha}}. \tag{11}$$

We refer to these investments as the *baseline ethical investment profile* $i(\omega)_k^* = \{h(\omega)_k^e, m(\omega)_k^e\}$. Plugging (10) and (11) into revenue from (6) gives equilibrium revenue generated by an ethical firm as

$$R(\omega)_k^e = A \alpha^{\frac{\alpha}{1-\alpha}} \left[\left(\frac{c_h}{\phi_k}\right)^{\beta} \left(\frac{c_m^e}{1-\phi_k}\right)^{1-\beta}\right]^{\frac{-\alpha}{1-\alpha}}. \tag{12}$$

Unethical Investments: We now turn to the non-cooperative investment game when the supplier has chosen the unethical technology. Demand is still given by equation (3), but the difference is that the indicator variable $I(\omega)$ may also take the value of zero. This is the case when the unethical firm does not mimic or if it faces an exogenous boycott in t_3. Mimicking involves setting the same price as the ethical firm. Therefore, the demand function becomes degenerate. When the unethical firm sets the ethical investment and price, $I(\omega) = 1$ and it gets full demand with probability $1-\gamma$. As soon as it deviates, we have $I(\omega) = 0$ and therefore zero demand.

An ethical firm faces a continuous demand function, leading to the continuous best response functions derived above. Consider the case that an unethical supplier would prefer mimicking over zero production. This is the only relevant case, as otherwise no supplier would choose unethical production in the first place. In this case the best response functions for the unethical firm are symmetric for the headquarter and the supplier and are given by

$$h(\omega)_k^u = \begin{cases} h(\omega)_k^e & \text{if } m(\omega)_k^u = m(\omega)_k^e \\ undetermined & \text{if } m(\omega)_k^u = 0 \\ 0 & otherwise \end{cases} \tag{13}$$

$$m(\omega)_k^u = \begin{cases} m(\omega)_k^e & \text{if } h(\omega)_k^u = h(\omega)_k^e \\ undetermined & \text{if } h(\omega)_k^u = 0 \\ 0 & otherwise. \end{cases} \tag{14}$$

Different to the ethical case, they take a value of zero for *any* investment of the other party deviating from the baseline ethical investment. The only point with positive investments of both parties is when they both set the baseline ethical investment.

While the best response functions are fundamentally different from the ones for the ethical firm, they share the Nash equilibria at zero-zero and the baseline ethical investments. In fact, they lead to the same equilibrium of the investment game. To see this, note that no party would ever find it optimal to choose an investment that is not on its best response function, as it would be strictly dominated by playing the best response. This implies that only two investments can occur for each party: zero or the baseline ethical investment. As in the case with ethical production we invoke the Pareto-dominance criterion so that the equilibrium with positive investment is the one that is played.

3.2.5 $t_1(a)$ (Un)ethical Technology Choice

We have seen how the non-cooperative investment decisions are taken for ethical and unethical firms in period $t_1(b)$. Based on this, we can now turn to period $t_1(a)$ analyzing the supplier's choice between the two technologies. In taking the technology decision, the supplier faces a trade-off between the cost savings implied by unethical production and the risk of losing its share of total revenues due to a consumer boycott.

First, consider the determinants of the expected revenues of the supplier. A (mimicking) unethical firm still faces a boycott with probability $1-\gamma$ so that expected revenues are given by $E\left[R(\omega)_k^u\right] = \gamma R(\omega)_k^e$. With a fraction $1-\phi_k$ going to the supplier and given the equilibrium $R(\omega)_k^e$ in equation (12), expected revenues of an unethical supplier are given by

$$(1-\phi_k)E\left[R(\omega)_k^u\right] = \gamma(1-\phi_k)A\alpha^{\frac{\alpha}{1-\alpha}}\left[\left(\frac{c_h}{\phi_k}\right)^{\beta}\left(\frac{c_m^e}{1-\phi_k}\right)^{1-\beta}\right]^{-\frac{\alpha}{1-\alpha}}. \tag{15}$$

The expected difference between ethical and unethical revenues of the supplier is

$$E\left[\Delta R_S\right] = \left(1-\phi_k\right)\left(R(\omega)_k^e - E\left[R(\omega)_k^u\right]\right).$$

This difference is always positive and reflects the fact that ethical firms have higher revenues in expectation, as they always face full demand. We refer to this difference as the *ethical revenue premium*. The supplier trades off its share of this ethical revenue premium against the cost savings of unethical production. The unit cost savings are determined by the scaling factor $\mu = \frac{c_m^u}{c_m^e}$ where $1-\mu \in (0,1)$ can be interpreted as the unit cost savings of unethical production which we refer to as the *unethical cost advantage*. Total cost savings of unethical production are given by $\Delta C = \left(c_m^e - c_m^u\right) m(\omega)_k^e$, with $m(\omega)_k^e$ given by equation (11).

In stage t_1, the organizational decision as well as the lump-sum transfer are fixed, as they are set in t_0. The supplier therefore takes the decision on unethical production by trading off $E\left[\Delta R_S\right]$ against ΔC. This decision can be described by a cutoff headquarter intensity β_S above which the supplier chooses the ethical technology and below which it produces unethically.

Lemma 1 *The headquarter intensity of a sector influences the technology choice of the supplier. Specifically, the supplier chooses unethical production when the headquarter intensity β is lower than*

$$\beta_S = 1 - \frac{1-\gamma}{\alpha\left(1-\mu\right)}. \tag{16}$$

The cutoff β_S (i) increases in the unethical cost advantage, $\frac{\partial \beta_S}{\partial(1-\mu)} > 0$; (ii) decreases in the probability of a boycott, $\frac{\partial \beta_S}{\partial(1-\gamma)} < 0$; (iii) and decreases in the mark-up, $\frac{\partial \beta_S}{\partial(1/\alpha)} < 0$.
Proof: *See the Appendix.*

Lemma 1 implies a direct link between headquarter/supplier intensity and (un)ethical production. Firms in sectors with a low headquarter (high supplier) intensity tend to implement the unethical technology, while ethical production is more likely in headquarter-intensive sectors. As the technology is not contractible, the supplier implements the type of technology that maximizes its individual expected profits. The choice between ethical and unethical production is driven by the trade-off between the *supplier's* total cost savings of unethical production and the *supplier's* expected loss of final revenue through a potential boycott. Cost savings of unethical production increase in the supplier's contribution to the production process. Therefore, suppliers in sectors with a low headquarter (high supplier) intensity (low β) tend to implement the unethical technology.

A stronger unethical cost advantage $1-\mu$ scales up total cost savings and makes unethical production attractive also for suppliers with lower levels of investments (i.e. in more headquarter-intensive industries). The supplier trades off these per-unit cost savings against the expected per-unit ethical revenue premium, which is determined by $1/\alpha$ and $1-\gamma$. The former represents the mark-up a firm charges over its marginal cost, representing the per-unit profit margin. The probability of facing a boycott $1-\gamma$ represents the risk of losing these profits when unethical production is chosen.

While the (un)ethical technology choice of the supplier depends on the headquarter intensity of production, the cost advantage of unethical production, the risk of a boycott and the mark-up (Lemma 1), it turns out to be *independent* of the organization of production. The reason is that the organization of

production has two countervailing effects on the attractiveness of unethical production for the supplier, which exactly offset one another. On the one hand, outsourcing scales up cost savings from unethical production through increased investments, on the other, it makes the supplier more vulnerable to a boycott by increasing its share of final revenues. This implies that the organization of production is an instrument for the headquarter to affect the supplier's investment only, and not its technology choice. Our model therefore remains very close to the setting in Antràs (2003) and allows us to focus on the prevalent question in this literature: how can one instrument (organization of production) be used to affect one variable (investments) under incomplete contracts? Our setting allows us to analyze how the potentially unethical technology choice of the supplier distorts the use of the instrument by the headquarter.

3.2.6 t_0: Optimal Organizational Structure and Transfer Payment

Transfer Payment Taking into account incomplete contracts, the investments in the manufacturing input and the equilibrium outcome of the ex-post bargaining, a supplier in a sector in which $\beta > \beta_S$ knows its private profits are going to amount to

$$\pi^e_{k,S} = (1 - \phi_k) R(\omega)^e_k - c^e_m m(\omega)^e_k \tag{17}$$

if it enters the match with the headquarter which has chosen organizational form $k \in \{O, V\}$. In the other case, in which a supplier knows it will choose unethical production and mimicking because $\beta < \beta_S$, it expects to earn

$$E\left[\pi^u_{k,S}\right] = \gamma (1 - \phi_k) R(\omega)^e_k - c^u_m m(\omega)^e_k \tag{18}$$

in case of a successful match. Because the headquarter faces a large number of potential suppliers competing perfectly for the opportunity to produce the final good with it, these expected private profits represent the maximum amount a supplier is willing to pay for this opportunity. The headquarter knows its own β and has decided the optimal organizational form $k \in \{O, V\}$. Given this decision and anticipating the technology choice of the supplier in t_1 the headquarter extracts

$$T_k = \begin{cases} \pi^e_{k,S} & \text{if } \beta > \beta_S \\ E\left[\pi^u_{k,S}\right] & \text{if } \beta < \beta_S. \end{cases} \tag{19}$$

Organizational Choice At the same time, the headquarter chooses between integration and outsourcing maximizing the total surplus of the match. Both decisions depend on the supplier's anticipated technology choice in stage t_1. For $\beta > \beta_S$, the headquarter anticipates ethical production by the supplier. In this case, the total surplus of the match is given by the sum of the two parties' private profits

$$\Pi^e_k = R(\omega)^e_k - c^e_m m(\omega)^e_k - c_h h(\omega)^e_k. \tag{20}$$

If $\beta < \beta_S$, the headquarter knows the supplier will choose the unethical technology and mimic an ethical firm in investments, quantities and prices. The total surplus of the match is then subject to the

uncertainty generated by the threat of a consumer boycott and is given by

$$E\left[\Pi_k^u\right] = \gamma R(\omega)_k^e - c_m^u m(\omega)_k^e - c_h h(\omega)_k^e. \tag{21}$$

In deciding the organizational form of the firm the headquarter compares the overall value of the relationship under outsourcing to the overall value under integration taking the technology choice of the supplier as given. Given ethical production by the supplier, the ratio of total profits under integration and total profits under outsourcing is given by

$$\Theta^e(\beta) = \left[\left(\frac{\phi_V}{\phi_O}\right)^{\beta}\left(\frac{1-\phi_V}{1-\phi_O}\right)^{1-\beta}\right]^{\frac{\alpha}{1-\alpha}} \frac{1-\alpha\left(1-\beta\right)+\phi_V\alpha\left[1-2\beta\right]}{1-\alpha\left(1-\beta\right)+\phi_O\alpha\left[1-2\beta\right]}.$$

The cutoff headquarter intensity above which the headquarter offers to the supplier a contract stipulating integration of the supplier and the transfer payment T_V given that it produces ethically ($\beta > \beta_S$) is implicitly defined by

$$\Theta^e(\beta_e) = 1. \tag{22}$$

Given unethical production by the supplier, the ratio of total expected profits is given by

$$\Theta^u(\beta) = \left[\left(\frac{\phi_V}{\phi_O}\right)^{\beta}\left(\frac{1-\phi_V}{1-\phi_O}\right)^{1-\beta}\right]^{\frac{\alpha}{1-\alpha}} \frac{\gamma-\alpha\left(1-\beta\right)\mu+\phi_V\alpha\left[\mu-\beta\left(1+\mu\right)\right]}{\gamma-\alpha\left(1-\beta\right)\mu+\phi_O\alpha\left[\mu-\beta\left(1+\mu\right)\right]}.$$

The cutoff headquarter intensity β^u above which the headquarter offers to the supplier a contract stipulating integration of the supplier and the transfer payment T_V given that it produces unethically ($\beta < \beta_S$) is implicitly defined by

$$\Theta^u(\beta_u) = 1. \tag{23}$$

The expression differs from $\Theta^e(\beta_e)$ in two respects. Because of unethical production there is now a threat of a boycott and second, the unethical cost advantage is exploited by the supplier. We summarize our result in the following subsection.

3.3 (Un)ethical Production, Headquarter Intensity and Ownership Structure

We can now combine the above insights on the implementation of the (un)ethical technology and the organizational choices of the firm conditional on technology to analyze the equilibrium of the model. Most notably, we are interested in the question of how the technology choice of the supplier interacts with the integration decision of the headquarter.

3.3.1 The Unethical Outsourcing Incentive

Based on equations (22) and (23), we can state the following proposition:

Proposition 1 *There exists a unique β_e below which the headquarter chooses outsourcing irrespective of the technology choice of the supplier. Integration is always chosen for headquarter intensities above β_u and it always holds that $\beta_e < \beta_u$. A sufficient condition for a unique interior solution $\beta_u \in (\beta_e, 1)$ to exist*

is given by $\gamma > \frac{4\phi_V}{3+\phi_V}$. *For any* $\beta \in (\beta_e, \beta_u)$ *the headquarter chooses integration if and only if the supplier produces ethically and chooses outsourcing if and only if unethical production is anticipated.*
Proof: *See the Appendix.*

The parameter condition $\gamma > \frac{4\phi_V}{3+\phi_V}$ is sufficient to ensure that $\beta_u < 1$ implying that both outsourcing and integration are chosen for some levels of headquarter intensity. Since we are interested in the interaction of unethical production with the organization of production, we focus on the cases in which both types of organizational form can emerge. However, $\beta_e < \beta_u$ regardless of whether the above condition holds.

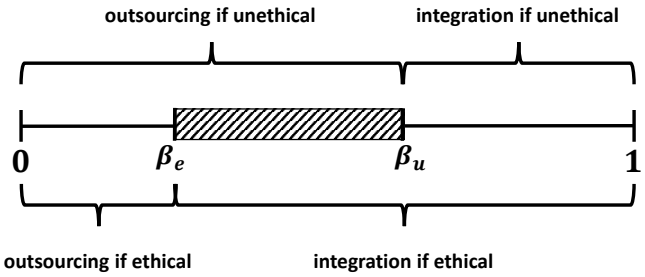

Figure 2: Unethical Production and the two Integration Cutoffs.

Figure 2 highlights the pattern described in Proposition 1. The axis shows the range of admissible headquarter intensities implying high supplier intensity on the left and high headquarter intensity on the right. The cutoff β_e is identical to the cutoff in Antràs (2003). It reflects the fact that the headquarter faces two underinvestment problems in period t_1 (the headquarter's and the supplier's). By (re)allocating residual rights of control, the organization of production is an instrument to alleviate the underinvestment of either the headquarter (through integration) or the supplier (through outsourcing).[33]

Unethical production by the supplier favors outsourcing: it reduces the unit costs of the manufacturing input so that the difference between the actual and the optimal investment of the (unethical) supplier increases. This aggravated underinvestment of the unethical supplier provides an additional incentive for outsourcing. We call this the *unethical outsourcing incentive*.

Due to the unethical outsourcing incentive, the cutoff separating outsourcing and integration changes when the headquarter expects unethical production. In the range between β_e and β_u, this additional outsourcing incentive leads the headquarter to *keep an unethical supplier at arm's length*, while it would opt for integrating an ethical supplier. This implies that for headquarter intensities between these two cutoffs, the organization of production *does* depend on the technology the supplier is anticipated to implement: the headquarter keeps an unethical supplier at arm's length *because* it is unethical.

3.3.2 Equilibrium Organization of Production

Define $\bar{\beta}$ as the headquarter intensity above which integration takes place. The actual value of $\bar{\beta}$ depends on how β_S relates to β_e and β_u. As $\beta_e < \beta_u$, three possible cases need to be considered.

[33]The mechanism is simply that integration and outsourcing imply different residual rights of control for the headquarter and the supplier. This changes the bargaining power and thereby the share of total revenue each party obtains. As a larger share of revenue increases the optimal investment, integration alleviates the headquarter's underinvestment while outsourcing alleviates the supplier's underinvestment.

Proposition 2 *(i) There exist three possible orderings of β_S, β_e and β_u that can arise in equilibrium: $\beta_e < \beta_S < \beta_u$ (Case 1); $\beta_e < \beta_u < \beta_S$ (Case 2) and $\beta_S < \beta_e < \beta_u$ (Case 3). (ii) The integration cutoff $\bar{\beta}$ differs across the three cases and is characterized by:*

$$\bar{\beta} = \begin{cases} \min\{\beta_S; \beta_u\} & \text{if } \beta_S > \beta_e \\ \beta_e & \text{otherwise.} \end{cases} \tag{24}$$

(iii) Unethical outsourcing and ethical integration are equilibrium outcomes in all three cases. Unethical integration and ethical outsourcing can occur in Cases 2 and 3, respectively.
Proof: *See the Appendix.*

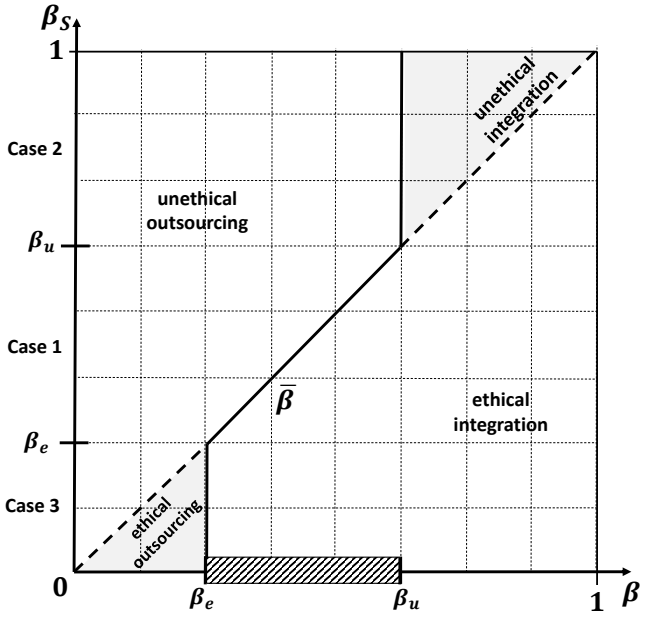

Figure 3: Equilibrium Organization of Production

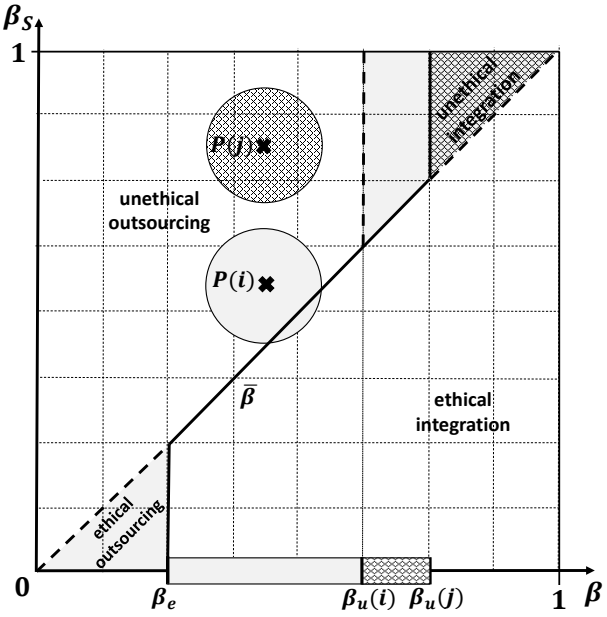

Figure 4: Sectoral Variation

Figure 3 illustrates the result stated in Proposition 2. The horizontal axis is taken from Figure 2. It plots the range of admissible headquarter intensities and highlights β_e and β_u. On the other axis, we plot the cutoff β_S, below which the supplier chooses unethical production and above which it implements the ethical technology. Along the diagonal we have $\beta = \beta_S$, i.e. we have unethical production above the diagonal and ethical production below.

The three cases referred to in Proposition 2 (i) differ in terms of the value of β_S in relation to the values of β_e and β_u. These cases are illustrated on the vertical axis. As for point (ii) in the above proposition, the solid line plots the *actual* integration cutoff $\bar{\beta}$ as defined in (24). Starting from the origin, and moving along the horizontal axis, it is given by $\bar{\beta} = \beta_e$ (Case 3), $\bar{\beta} = \beta_S$ (Case 1) and $\bar{\beta} = \beta_u$ (Case 2). Outsourcing thus takes place to the left of the $\bar{\beta}$-line and integration to the right of it. The intuition for the shape of this relationship is as follows: for $\beta < \beta_e$, outsourcing is chosen even under ethical production. Therefore, no integration occurs no matter how high or low β_S. For the range of $\beta_e < \beta < \beta_u$, a strong unethical cost advantage *can* induce outsourcing. In this range, outsourcing is always linked to unethical production and integrated suppliers always produce ethically. Finally, for very high values of the headquarter intensity

($\beta > \beta_u$), the underinvestment of the headquarter is so severe that integration is optimal for any value of β_S. Moreover, the graph highlights point (*iii*) of the above proposition: the diagonal ($\beta = \beta_S$) separates ethical and unethical production. As $\bar{\beta}$ (which separates outsourcing and integration) does not coincide with the diagonal, unethical integration and ethical outsourcing can occur in equilibrium in Cases 2 and 3, respectively.

Figure 3 also nicely illustrates the difference of our model to the Antràs (2003) benchmark. Our model nests the Antràs (2003) model when unethical production is never chosen in equilibrium. This case occurs for $\beta_S \leq 0$ so that β_e is always the relevant integration cutoff. Because β_e is identical to the cutoff from Antràs (2003), the two models then coincide.

3.3.3 Link Between Theory and Empirics

We have now identified the determinants of the integration vs. outsourcing decision in our model and have seen how the pattern changes compared to Antràs (2003). In particular we have seen how the implementation of the unethical technology alters the predictions of the model due to the unethical outsourcing incentive. We can now proceed to determine whether the model is in line with the examples in the Introduction and the empirical analysis in Section 2. Those results suggest a positive relationship between the incentive for unethical production and outsourcing. The main result in Section 2 was that - controlling for the standard determinants of integration - a higher environmental cost savings potential (ECSP) is associated with more outsourcing and the more so, the lower the level of the environmental policy stringency index (EPSI) in the source country. The unethical cost advantage $1-\mu$ is the theoretical counterpart of the combination of ECSP and EPSI and thus combines industry-country-specific information about the benefits of unethical production. To determine whether our theory can rationalize the pattern in the data, we use our model to investigate the effect of the unethical cost advantage $1-\mu$ on the actual integration cutoff $\bar{\beta}$. We show that the theory indeed predicts the empirical pattern: (controlling for the headquarter intensity of production) a higher unethical cost advantage is associated with more international outsourcing.

Equation (24) allows us to determine the impact of the potential cost savings of unethical production on the integration vs. outsourcing decision in our model. Our model indeed delivers an unambiguous prediction:

Proposition 3 *The outsourcing cutoff is weakly increasing in the unethical cost advantage, i.e.* $\frac{\partial \bar{\beta}}{\partial(1-\mu)} \geq 0$.

Proof: *See the Appendix.*

We can see from Proposition 3 that the integration cutoff is weakly increasing in the unethical cost advantage given by $1-\mu$. For $\beta \in (\beta_e, \beta_u)$, the relevant cutoff is β_S, which, by Lemma 1, increases in $1-\mu$. For $\beta \in (\beta_e, \beta_u)$, a higher $1-\mu$ therefore implies that there is a larger range of sectors whose headquarter intensity is below $\bar{\beta}$. In Figure 4, the area to the left of the $\bar{\beta}$-line increases as β_S rises. For $\beta > \beta_u$, we have $\bar{\beta} = \beta_u$. We have seen in the proof of Proposition 1 that β_u is also increasing in $1-\mu$. An increase in $1-\mu$ therefore also moves the $\bar{\beta}$-line to the right at $\bar{\beta} = \beta_u$, also increasing the range of headquarter intensities located to the left of $\bar{\beta}$. This implies that for all headquarter intensities $\beta > \beta_e$,

the range of sectors in which firms choose outsourcing increases as $1-\mu$ rises. For $\beta < \beta_e$ outsourcing is chosen for any value of β_S.

A close link to the empirical analysis requires some more arguments. In the empirical analysis in Section 2, the *share* of intrafirm trade in a sector is used as the dependent variable. It is interpreted as a measure of the prevalence of integration in a sector. Our theoretical analysis concerns a single headquarter-supplier match in a single sector.

In Antràs (2003), only the headquarter intensity β varies across sectors. In order to link our theory to the empirical analysis in Section 2 we need to allow for $1-\mu$ to also vary across sectors: different industries feature different potential cost savings from unethical production.

Taking our model at face value, we should observe either *only* integration or *only* outsourcing in each sector. In the empirical analysis, however, we use the *continuous* share of intrafirm trade in each industry as our dependent variable. The working paper version of Antràs (2003) provides an argument why even in a model with homogeneous firms this type of empirical analysis is appropriate to test the model's theoretical predictions.[34] In what follows, we generalize his argument to many sectors and show that our explorative empirical analysis in Section 2 directly speaks to our theoretical results.

In Figure 4, a sector i is characterized by a headquarter intensity $\beta(i)$ and a cutoff $\beta_S(i)$, as we consider $1-\mu$ sector-specific. As any admissible combination of values of β and β_S is represented by one point in Figure 4, there is a single point P_i representing sector i in the graph (of course also other sectors may share the same point with sector i). Assume now that there is a large number of firms in each sector and that firms are uniformly distributed across the space in Figure 4. Further, assume that the statistician is unable to correctly allocate all firms to their sectors and that for firms in any sector $j \neq i$ there is a positive probability of being misallocated to sector i and that this probability increases in the similarity between sector i and j. By similarity we mean similarity in terms of β and μ. In the case of Figure 4, this translates into similarity in terms of β and β_S, as the latter is a function of μ. For the sake of the argument, define the similarity between sector i and sector j as $s_{ij} = 1 - \big(\epsilon(\beta(i)-\beta(j))^2 + (1-\epsilon)(\beta_S(i)-\beta_S(j))^2\big)$, where $\epsilon \in (0,1)$ represents arbitrary weights. With this metric at hand, compared to sector j, all sectors can be classified as more, equally, or less similar to sector i. Moreover, this allows to draw iso-similarity lines, as illustrated in Figure 4.

Define i^e as the empirical counterpart of sector i constructed by the statistician. It contains correctly allocated firms from sector i but - by the law of large numbers - also misallocated firms from all other sectors. We can use iso-similarity lines for different levels of similarity to illustrate the empirical sector i^e. Consider for example the point P_i in Figure 4. It is located above the diagonal ($\beta_S(i) > \beta(i)$) and $\beta(i) \in (\beta_e(i), \beta_u(i))$. Therefore, all firms in the *actual* sector i opt for unethical outsourcing. When we consider iso-similarity lines for decreasing levels of similarity, at some point they will cross the $\bar{\beta}$-line. This implies that in the construction of the empirical sector i^e, some firms choosing ethical integration are (mis)allocated to sector i^e.

This has two important implications. First, sector i^e contains some firms choosing ethical integration (and in fact, also some firms choosing ethical outsourcing and unethical integration). The empirically observed share of intrafirm trade flows in total trade flows of sector i^e (the empirical measure of the prevalence of integration in sector i^e) will therefore not be equal to zero, even though none of the firms

[34] See NBER Working Paper 9740, p.37.

actually belonging to i choose integration. Second, the probability of a firm being misallocated into sector i^e decreases in the similarity-distance of that firm from P_i. Therefore, everything else equal, the empirical counterpart j^e of an actual sector j with a stronger unethical cost advantage ($1-\mu(j) > 1-\mu(i)$) characterized by a point P_j above point P_i should exhibit a *smaller* (though still positive) fraction of intrafirm trade. A stronger unethical cost advantage also increases the cutoff β_u, which reinforces the above effect: as illustrated in Figure 4, for the sector with the stronger unethical cost advantage the vertical part of the $\bar{\beta}$-line at $\beta = \beta_u$ is located further to the right and enlarges the area in which unethical outsourcing is chosen.

Proposition 3 therefore speaks directly to the empirical analysis in Section 2: comparing two sectors i and j that have the same headquarter intensity β but j has a stronger unethical cost advantage, we have $\beta_S(j) > \beta_S(i)$ and P_j therefore lies vertically above P_i in Figure 4. This is in line with the core empirical finding: controlling for headquarter intensity $\beta(i)$, the model suggests that we should indeed observe a larger share of outsourcing in sectors with a high unethical cost advantage, which is what we find in the data.[35]

3.3.4 Headquarter's Perspective on Ethical Production: Aspirations and Reality

Before we proceed to analyzing the role NGOs may play for the link between the boycott and prices, output and investments in Section 4, we now highlight an interesting tension that can arise between the headquarter's aspirations and actions regarding (un)ethical production. Consider a headquarter that states that it would like to source its products ethically but then incentivizes its suppliers to expand unethical production. An external observer may interpret this as evidence of a dishonest attempt at greenwashing or - simply put: a lie - by the firm. Our model, however, implies that this can actually be an equilibrium outcome.

For this to arise, we need to be in a situation where the headquarter chooses outsourcing if and only if unethical production is anticipated, i.e. $\beta \in (\beta_e, \beta_u)$. We have seen in the discussion of Proposition 1 that this is the range in which the unethical outsourcing incentive leads headquarters to choose outsourcing, while they would choose integration in its absence. This corresponds to the shaded range of headquarter intensities in Figure 2.

Moreover, as we have seen in the proof of Lemma 1, the supplier trades off the cost savings of unethical production only against its *own* fraction $1-\phi_k$ of the expected loss in revenues from unethical production. This delivers the cutoff β_S in Lemma 1, above which the supplier chooses ethical production. Different to the supplier, the headquarter can extract, and therefore seeks to maximize, the *full expected profits* of the match. If the headquarter were able to choose the technology used in production (which it cannot due to incomplete contracts on technology), its preferred technology may differ from the supplier's choice.

In analogy to β_S for the supplier, define the technology cutoffs $\beta_{H,k}$ with $k \in \{V, O\}$ as the cutoff headquarter intensities above which - conditional on the organization of production - the headquarter would choose ethical production. We can then state the following proposition.

[35] Note that this is the case for any admissible point in the graph, no matter where the starting point P_i lies: a vertical upward shift decreases the distance to the sectors in the upper-left area of the graph (doing outsourcing) and increases the distance to the lower-right area. This implies that sectors choosing unethical outsourcing get a stronger weight in the empirical sector i^e.

Proposition 4 *The technology cutoffs maximizing full expected profits satisfy $\beta_{H,V} < \beta_{H,O} < \beta_S$. There is a non-empty set of headquarter intensities that satisfy $\beta \in (\beta_{H,O}, \beta_S) \wedge \beta \in (\beta_e, \beta_u)$. That is, the headquarter would oblige the supplier to produce ethically if it could, but, as it cannot, chooses outsourcing in order to expand unethical production.*
Proof: *See the Appendix.*

Figure 5: Aspirations vs. Reality

The solid area in Figure 5 illustrates this result. Between $\beta_{H,O}$ and β_S, the supplier implements the unethical technology, while the headquarter would prefer the ethical one. At the same time the range of headquarter intensities in question lies between β_e and β_u, which implies - due to the unethical outsourcing incentive - that anticipation of unethical production by the supplier leads the headquarter to choose outsourcing in order to fully exploit the cost advantage of unethical production. This result arises wherever the intervals (β_e, β_u) and $(\beta_{H,O}, \beta_S)$ overlap.

Therefore, due to incomplete contracts and the unethical outsourcing incentive, a headquarter stating truthfully that it would like to source its products ethically (if it could), but at the same time incentivizing an expansion of unethical production can be an equilibrium outcome of our model.

4 Extension: NGOs, Firms, and Consumer Boycotts

In this section, we introduce elements of Private Politics to explicitly model the link between firm choices and consumer boycotts. Our extension features private information on technology and an advocacy NGO investigating firms. We take a clear stand on how consumer boycotts emerge and how unethical production affects the risk of facing a boycott. While in the baseline model we imposed that failure of an unethical firm to pool with ethical firms leads to a boycott with certainty, we derive this here from the expectation formation of the NGO and firms' equilibrium response to it. We show that the qualitative results of the baseline model as well as the relation to the empirics we derive from it continue to hold in this extension of the model.

4.1 Co-Existence of Ethical and Unethical Firms

We assume that the type of technology used is private information of the headquarter-supplier match. With the technology used in production being a credence attribute, i.e. consumers care about it but cannot infer it from the final product, private information allows unethical producers to pass as ethical firms and avoid a boycott.

In the baseline model, either all firms in a sector choose the ethical technology or all choose the unethical technology. This is a very stylized pattern that directly stems from the fact that all firms in a sector are identical. In a sector in which all firms implement the unethical technology, mimicking would not make sense, as there are no ethical firms to pool with. We therefore assume that only a fraction κ

of suppliers in each sector is able to use the unethical technology. Because of this, in equilibrium there will be at least a fraction $1-\kappa$ of firms that produce ethically. In period t_0, when the headquarter offers the transfer payment to the supplier and decides the organizational form of the firm, neither party knows whether unethical production will be possible. This is only revealed at the next stage just before investment decisions are taken and the (un)ethical technology choice is made.[36] This assumption implies that the organizational choice of the firm does not contain information on the type of the firm: when it is taken, the headquarter does not know whether the unethical technology will be available in period $t_1(a)$.

4.2 NGO Investigations and Consumer Boycotts

We assume that there is an NGO that is able to investigate firms and to organize such boycotts. As the focus of this paper remains on the international organization of production, we keep the modeling of the NGO relatively stylized.[37]

While technology cannot be directly observed from outside the match, output and prices can be used to infer investment quantities. We assume that the NGO is sophisticated enough to determine the optimal choices of an ethical firm in a given sector. It then potentially faces two types of firms. First, firms that deviate from these choices and are therefore openly unethical. In this case, identifying the firm as unethical is costless for the NGO (we will see below that this is not an assumption, but an outcome of the extended model). Second, a group of *seemingly ethical firms* that are all identical in terms of observables, but which contains ethical and (mimicking) unethical firms. When unethical production is profitable in expectation, the NGO knows that a fraction κ of firms among the seemingly ethical firms are in fact unethical.

As for the group of seemingly ethical firms, we assume the NGO's investigation to be costly. And indeed, in reality, the acquisition of verifiable information on pollution and working conditions and the link to final consumer brands is a costly and possibly dangerous (and illegal) activity in many countries. One example is the Detox campaign by Greenpeace addressing, among other things, the toxic water pollution of the Pearl and Yangtze River Deltas (Greenpeace, 2011) and the Qiantang River (Greenpeace, 2012) in China by local textile and apparel producers. According to Greenpeace, a year-long investigation into production practices and buyer-seller linkages preceded its campaign to push a large number of top labels in the apparel industry to 'detox' their supply chain. Another well-mediatized example of the dangers of investigating working conditions in countries like China is the case of a labor activist being arrested for trying to document poor working conditions in a factory producing shoes for Ivanka Trump's brand in southern China (New York Times, 2017).

The result of an investigation is that the NGO learns about the technology used by a given firm. If the firm is of the ethical type, no boycott is started. If the firm is unethical, the NGO can start a campaign, i.e. it recommends to consumers to boycott the firm. While investigating the technology implemented by the supplier is costly for the NGO, we assume for simplicity that triggering the consumer boycott is

[36] One way to think about this is as follows. Ex-ante the supplier knows that there is some probability κ that it can e.g. bribe government officials to turn a blind eye on toxic waste disposal into a river or on the violation of work safety standards. If this is actually possible in the individual case, only turns out after the match is formed and some investments are made.

[37] Different to e.g. Krautheim and Verdier (2016) or Aldashev and Verdier (2009) we do not intend to contribute to a better understanding of the endogenous emergence of NGOs, interactions with donors, the trade-offs shaping the fundraising process or the optimal allocation of funds across firms or sectors.

costless.

As outlined in the Introduction, consumers consistently report NGOs to be the most trusted institutions. Our modeling reflects this relationship of trust in that consumers do not request the type of evidence from the NGO that would be required in a court of justice. They take their boycott decision based on the subjective judgment of the NGO. To illustrate the point, consider the following example: an advocacy NGO sends an agent into a polluting production plant in China. The agent reports (and provides photographs showing) that workers in the plant stitched a well-known label into the shirts produced there. For the NGO this may be sufficient evidence to call for a boycott – and given the trust in the NGO – this may also be sufficient for consumers not to buy that label anymore. But this would clearly not be sufficient evidence (one witness, some photos - even obtained in an illegal private undercover operation) to lead to a conviction in a court of justice, e.g. if the owner of the label sued its supplier for unethical practices. It is important to note that this implies that NGOs may find out about unethical production and consumers may engage in a boycott while the implemented technology remains - legally - unverifiable and therefore not contractible between the headquarter and the supplier.[38]

As investigations are costly, the fraction of firms $1-\gamma$ the NGO can monitor depends on the funds F it can raise. To organize ideas, we assume that this relation is determined by $1-\gamma = \Psi(F)$, where $\Psi(F)$ is strictly increasing in F. Also here, we keep the modeling very stylized and simply take the funds F as exogenous.

These are the extensions and refinements we make in order to model the endogenous occurence of consumer boycotts. All other events in the different periods are just as in the baseline model. We will next discuss the belief formation of the NGO and the informational content of the firms' choices.

4.3 NGO Beliefs and Investigations

The non-cooperative investment game results for each firm in an observable investment profile $i(\theta) = \{h(\theta), m(\theta)\}$ with $h(\theta) \geq 0$ and $m(\theta) \geq 0$.[39] In period t_3, the NGO picks an action $s_i \in \{0,1\}$ which is to initiate an investigation on a firm with investment profile i or not.[40]

[38]To reiterate on this important point, note the following: (1) Courts of justice are the institutions that enforce contracts. (2) For the property-rights theory of the firm to work, is has to be assumed that investments (here also technology) cannot be verified *in a way that would lead to the conviction of one of the parties in a court trial.* (3) It is therefore consistent with the property-rights theory of the firm to assume that a third party can gather some information on the investments/technology and establish probable cause (e.g. of unethical production) as long as this is not sufficient evidence to hold in court. We argue that this assumption is plausible as courts operate under the 'in dubio pro reo' principle and apply high formal standards to what can count as evidence, while for consumers convincing circumstantial evidence evaluated by a *trusted* institution is sufficient.

[39]Note two things. First, we speak of an observable investment profile, however, as outlined in Footnote 30, investments can be inferred from optimal quantities and prices, so they do not necessarily have to be observable. Second, for ease of exposition we suppress the organizational subscript k and the variety index ω where possible. It is well understood that the strategies are chosen and decision are made conditional on outsourcing or vertical integration chosen by the headquarter at an earlier point in the game.

[40]When the investments are interpreted by the NGO as containing information on the implemented technology, one could think that there is room for strategic signaling when setting investments. This would place us in the context of a signaling game similar to the one in Krautheim and Verdier (2016). The core idea of the signaling literature in economics (Spence 1973, 1974) is that an agent of a 'high' type may deviate from an otherwise optimal action for the sole purpose to differentiate itself from a 'low' type which would otherwise pool with the 'high' type in terms of observables. This requires that all parties understand that an action is taken on purpose in order to signal one's type. The obvious difference to our setting is that investments - the decision that contains information about the type of the firm - are set non-cooperatively. With incomplete contracts headquarter and supplier can not even coordinate on the profit maximizing investment, let alone an investment in order to signal their type.

The NGO has a belief function $\eta(\theta \mid i)$. Conditional on observing some investment profile i, it assigns a probability of $\eta(\theta \mid i)$ to the firm being of type θ. If $\eta(\theta = u \mid i) = 1$, the NGO immediately starts an investigation.

Proposition 5 *In the extended model, (i) ethical firms are indifferent to NGO investigations and therefore set their investments independently of NGO beliefs; (ii) unethical firms face an NGO investigation with certainty unless they mimic (i.e., set the same investment as) ethical firms. If unethical firms mimic ethical firms, their probability of being investigated is reduced to* $1 - \gamma < 1$.
Proof: *In the text.*

The expectations of the NGO follow Bayes' Law implying the following belief function

$$\eta(\theta = e \mid i) = \frac{Pr(i \mid \theta = e)\ Pr(\theta = e)}{Pr(i \mid \theta = e)\ Pr(\theta = e) + Pr(i \mid \theta = u)\ Pr(\theta = u)}. \tag{25}$$

Note that ethical firms are indifferent to being investigated: they always get full demand in period t_5, as they never face a boycott. Denote by $\tilde{i}$ the investment profile of an ethical firm resulting from the non-cooperative investment game. Even if it could, an ethical firm would never adjust $\tilde{i}$ to accord with an arbitrary belief of the NGO, as this only affects the probability of being investigated, which has no effect on the firm.

We therefore have $Pr(\tilde{i} \mid \theta = e) = 1$ and $Pr(\bar{i} \mid \theta = e) = 0$ for any $\bar{i} \neq \tilde{i}$. Therefore, $\tilde{i}$ is the only investment profile for which the NGO assigns a positive probability to ethical production: $\eta(\theta = e \mid \tilde{i}) > 0$ and $\eta(\theta = e \mid \bar{i}) = 0$ for any $\bar{i} \neq \tilde{i}$. Any other investment profile triggers an immediate investigation by the NGO.

The NGO can compute if in a given sector firms have an incentive to be unethical. When unethical firms in that sector pool with ethical firms by setting $\tilde{i}$, they form a group of *seemingly ethical firms* for which investigation is costly for the NGO. As in this case $\eta(\theta = e \mid \tilde{i}) < 1$, the NGO trivially maximizes its objective of starting a boycott against the largest possible number of unethical firms by spending its whole budget on investigations of firms in the seemingly ethical group (and then start costless boycotts against all identified unethical firms).[41]

[41] One may think that the fact that the NGO interprets the investment levels as containing information of the type of the firm can only work in a context of homogeneous firms; as with heterogeneous firms, each firm has a different 'ethical' output level. Heterogeneous levels of production, investments and prices, however, are not a problem *per se*. Additional complexity only arises if the underlying driver of heterogeneity is unobservable to the NGO. The most straightforward way of introducing firm heterogeneity would therefore be to assume firm heterogeneity in quality, which has become a standard modeling approach in the literature on trade and quality (see e.g. Baldwin and Harrigan, 2011 and references therein). Being part of the utility function of the consumer, quality can hardly be private information of the firm, so that the intuition of our model extension goes through also with heterogeneous firms. Some comments on heterogeneity in quality vs. productivity are in order. First, already Melitz (2003) points out that modeling heterogeneity in quality and productivity is isomorphic (Footnote 7, p. 1699). Second, models with heterogeneous quality imply that larger firms set higher prices, a stylized fact that heterogeneity in productivity cannot capture (see, e.g. Kugler and Verhoogen, 2012). Third, the empirical measure of 'productivity' is a residual which could equally well be attributed to 'quality'. But even the introduction of heterogeneity in productivity (combined with productivity being private information of the firm) would not be inconsistent with our modeling. In this case, the NGO would know the distribution of 'ethical' prices, investments and output levels and the actual distribution. For those e.g. prices where the actual distribution deviates most from the 'ethical' distribution, the NGO may choose stronger levels of investigation. We do not expect this additional layer of complexity to add substantially to addressing the research questions of this paper.

4.4 Non-Cooperative Investments with Degenerate Demand

We have seen above that unethical firms can only generate positive demand (in expectation) by investing $\tilde{i}$. For this investment the firm faces full demand if it arrives at stage t_5 without a boycott.

Lemma 2 *The equilibrium investment profile $\tilde{i}$ of an ethical firm is characterized by the same expressions, i.e. equations* (10) *and* (11)*, as the equilibrium profile i^* in the baseline model.*
Proof: *This directly follows from the fact that the optimal choices of the headquarter and the supplier in a match that only has the ethical technology available (or in a sector where all firms endogenously choose ethical production), is unaffected by any element of the model extension.*

It remains to be shown that $\tilde{i} = i^*$ is the equilibrium outcome of the non-cooperative investment game also for an unethical firm. Clearly, it is a Nash equilibrium of the investment game if it yields positive profits in expectation, as any deviation from it would lead to zero demand. As in the Antràs (2003) model, zero-zero is a Nash equilibrium that is ruled out by the Pareto dominance assumption.

Consider the case of an unethical firm. It follows from Equations (13) and (14) that the best response to any investment level other than $i_k^* = \{h_k^e, m_k^e\}$, with $k \in \{V, O\}$ is zero for both parties, as any deviation from i_k^* leads to an investigation by the NGO resulting in a boycott with zero demand. No party would ever find it optimal to choose an investment that is not on its best response function, as it would be strictly dominated by playing the best response. We can therefore state the following proposition.

Proposition 6 *In the extended model, unethical firms mimic ethical firms, i.e. the equilibrium investment profile of an unethical firm is identical to the equilibrium investment profile of an ethical firm.*
Proof: *In the text.*

Using the results of this section, we show in the Online Appendix that the extended model delivers the same qualitative results as the baseline model. Expressions only differ as they now also contain the fraction $1 - \kappa$ of firms that cannot use the unethical technology, which we introduced for consistency in the extension.

5 Conclusion

Over the last decades the internationalization of the value chain has allowed firms to exploit cross-country differences in environmental and labor regulation (and enforcement) in ways that are frequently criticized as 'unethical' and have triggered a large number of NGO campaigns and consumer boycotts. In this paper we addressed the question of how potential 'unethical' cost savings on the one hand and the threat to reputation and sales on the other interact with the international organization of production.

We identified a large number of examples of NGO campaigns and consumer boycotts that seem to suggest a link between international outsourcing and unethical production. To investigate this aspect more systematically, we carried out an explorative empirical analysis. As a measure of vertical integration at the industry level we used the share of intrafirm imports in total U.S. imports. We correlated this variable with a measure of the cost advantage of unethical production at the industry level, providing evidence that a high cost advantage of unethical production in an industry indeed favors international outsourcing

as opposed to vertical integration. We found this effect to be strongest for sourcing from countries with a low level of environmental regulation.

In our baseline model, we introduced a cost-saving 'unethical' technology and consumer boycotts into a standard property-rights model of the international organization of production. Like in Antràs (2003), we found that a high supplier intensity favors international outsourcing. The first central result of our analysis was that with incomplete contracts on technology, suppliers in sectors with a high supplier intensity tend to implement the unethical technology. Our second central result was that we identified a new driver of the integration vs. outsourcing decision: the *unethical outsourcing incentive*. It generates a range of supplier intensities for which the headquarter chooses outsourcing if and only if it anticipates unethical production, i.e. it keeps the supplier at arm's length just because it is unethical.

We then analyzed how incomplete contracts on technology and the unethical outsourcing incentive affect the equilibrium organization of the firm. We found that a headquarter anticipating unethical production by its supplier tends to opt for outsourcing and that this bias towards outsourcing increases in the cost advantage of unethical production. This implies that, everything else equal, more outsourcing arises in sectors with a stronger cost advantage of unethical production. We argued that the implications of the theory therefore directly relate to the findings of the explorative empirical analysis: controlling for other sector characteristics (especially the supplier intensity of production), firms in sectors with a high potential for unethical cost savings tend to keep suppliers at arm's length. Moreover, we showed that it is possible that the headquarter would prefer *ethical* production (but cannot impose it due to incomplete contracts) and optimally responds with incentivizing an expansion of *unethical* production through outsourcing.

To focus on the implications of unethical production for the international organization of production, in the baseline model, we imposed that any deviation from investments, quantities or prices of an ethical firm immediately triggers a consumer boycott. We also analyzed an extension of the model in which the link between a deviation from the ethical observables and a boycott emerges from the expectation formation of an NGO, which can conduct investigations and organize consumer boycotts, as well as firms' equilibrium response to it. We found that all results from the baseline model hold qualitatively.

Appendix

A.1 Proof of Lemma 1

A.1.1 Derivation of the supplier cutoff

For unethical production to be preferred, we need the total cost savings from unethical production ΔC to be larger than the expected ethical revenue premium $E[\Delta R_S]$.

$$\begin{aligned}
\Delta C &> E[\Delta R_S] \\
(c_m^e - c_m^u)\, m(\omega)_k^e &> (1-\phi_k)\,(R(\omega)_k^e - E[R(\omega)_k^u]) \\
(c_m^e - c_m^u)\, m(\omega)_k^e &> (1-\phi_k)\,(1-\gamma)\, R(\omega)_k^e \\
(c_m^e - c_m^u)\,(1-\beta)\, A\alpha^{\frac{1}{1-\alpha}}\, \frac{1-\phi_k}{c_m^e} \left[\left(\frac{c_h}{\phi_k}\right)^{\beta}\left(\frac{c_m^e}{1-\phi_k}\right)^{1-\beta}\right]^{-\frac{\alpha}{1-\alpha}} & \\
> (1-\gamma)\,(1-\phi_k)\, A\alpha^{\frac{\alpha}{1-\alpha}} \left[\left(\frac{c_h}{\phi_k}\right)^{\beta}\left(\frac{c_m^e}{1-\phi_k}\right)^{1-\beta}\right]^{-\frac{\alpha}{1-\alpha}} & \\
\frac{c_m^e - c_m^u}{c_m^e}\,(1-\beta)\,\alpha &> 1-\gamma
\end{aligned}$$

Solving for β using the fact that $c_m^u = \mu c_m^e$ gives that when

$$\beta < \beta_S = 1 - \frac{1-\gamma}{(1-\mu)\,\alpha}, \tag{A.1}$$

the supplier will prefer unethical production.

Comparative statics Differentiating w.r.t. $1-\mu$, $1-\gamma$, and $\frac{1}{\alpha}$ delivers

$$\frac{\partial \beta_S}{\partial (1-\mu)} = \frac{1-\gamma}{\alpha\,(1-\mu)^2} > 0. \tag{A.2}$$

$$\frac{\partial \beta_S}{\partial (1-\gamma)} = -\frac{1}{\alpha\,(1-\mu)} < 0. \tag{A.3}$$

$$\frac{\partial \beta_S}{\partial \frac{1}{\alpha}} = -\frac{1-\gamma}{(1-\mu)} < 0. \tag{A.4}$$

A.2 Proof of Proposition 1

The cutoff β_l is the value of β that solves

$$\Theta^l(\beta_l) = \left[\left(\frac{\phi_V}{\phi_O}\right)^{\beta_l}\left(\frac{1-\phi_V}{1-\phi_O}\right)^{1-\beta_l}\right]^{\frac{\alpha}{1-\alpha}} \frac{\gamma - \alpha\,(1-\beta_l)\,\mu + \phi_V \alpha\,[\mu - \beta_l\,(1+\mu)]}{\gamma - \alpha\,(1-\beta_l)\,\mu + \phi_O \alpha\,[\mu - \beta_l\,(1+\mu)]} = 1 \tag{A.5}$$

with $\gamma, \mu \in (0,1)$ delivering the unethical cutoff β_u. In the corner case of $\gamma = \mu = 1$, the β that solves the equation is β_e.

A.2.1 Existence

To show existence of the two cutoffs, we will derive conditions under which the corner cases $\Theta^l(\beta = 1) > 1$ and $\Theta^l(\beta = 0) < 1$ are true, implying that there exists some β_e for which $\Theta^e(\beta_e) = 1$ and some β_u for which $\Theta^u(\beta_u) = 1$.

Case 1: $\beta = 0$ $\Theta^l(\beta)$ reduces to

$$\left(\frac{1-\phi_V}{1-\phi_O}\right)^{\frac{\alpha}{1-\alpha}} \left[\frac{\gamma - \alpha\mu\,(1-\phi_V)}{\gamma - \alpha\mu\,(1-\phi_O)}\right]. \tag{A.5'}$$

Case 2: $\beta = 1$ $\Theta^l(\beta)$ becomes

$$\left(\frac{\phi_V}{\phi_O}\right)^{\frac{\alpha}{1-\alpha}} \frac{\gamma - \phi_V\alpha}{\gamma - \phi_O\alpha}. \tag{A.5"}$$

Here, again, $\gamma, \mu \in (0,1)$ deliver Θ^u and $\gamma = \mu = 1$ deliver Θ^e.

Numerator and denominator of each of the two cases differ only in the value of ϕ_k. Substituting x for $1 - \phi_k$ in (A.5') and for ϕ_k in (A.5") and recalling that $\frac{1}{2} < \phi_k < 1$, the two cases only differ in the value of μ. Numerator and denominator of any of the two cases can be expressed in general form as

$$x^{\frac{\alpha}{1-\alpha}} (\gamma - \alpha\mu x)\,. \tag{A.6}$$

Because $\phi_V > \phi_O$ (and thus $1 - \phi_V < 1 - \phi_O$), conditions that ensure that equation (A.6) has a positive slope in x also ensure that $\Theta^l(\beta = 0) < 1$ and $\Theta^l(\beta = 1) > 1$.

$$\frac{\partial}{\partial x} x^{\frac{\alpha}{1-\alpha}} (\gamma - \alpha\mu x) = \frac{\alpha}{1-\alpha} x^{\frac{\alpha}{1-\alpha}} \left(\frac{\gamma}{x} - \mu\right)$$

Because $x \in (0,1)$ and $\alpha \in (0,1)$, the last factor determines the sign of the derivative. We must cover four cases, each of Cases 1 and 2 from above for ethical ($\gamma = \mu = 1$) and unethical production, i.e. with $\gamma, \mu \in (0,1)$. For ethical production, we require $(1/x) - 1 \overset{!}{>} 0$ for $\beta = 0$ and $\beta = 1$. For unethical production, we require $(\gamma/x) - \mu \overset{!}{>} 0$ for $\beta = 0$ and $(\gamma/x) - 1 \overset{!}{>} 0$ for $\beta = 1$. For ethical production, the condition always holds because $\frac{1}{x} > 1$ in both cases. To ensure existence of β_u, both conditions under unethical production must hold, i.e. we must have $\gamma > \mu\,(1 - \phi_O)$ and $\gamma > \phi_V$. As $\gamma > \phi_V$ is the stricter condition, it is also a sufficient condition for existence.

Therefore, with ethical production, $\Theta^e(\beta = 1) > 1$ and $\Theta^e(\beta = 0) < 1$, and hence, β^e exists. With unethical production, if $\gamma > \phi_V$, then $\Theta^u(\beta = 1) > 1$ and $\Theta^u(\beta = 0) < 1$, therefore β^u exists. QED.

A.2.2 Uniqueness

To establish uniqueness, we show under which conditions the derivative of $\Theta^l(\beta)$ with respect to β is larger than zero for all $\beta \in [0,1]$. The proof follows the structure of Appendix 2 in Antràs (2003).

Recall that $\phi_V = \phi_O + \delta^\alpha\,(1 - \phi_O)$, where δ is the share of the intermediate the headquarter can continue to use in an integrated firm in case bargaining breaks down. Using this relationship, $\Theta^l(\beta)$ can

be written as

$$\Theta^{l}(\beta)=\underbrace{\left[1+\frac{\delta^{\alpha}}{\phi_{O}\left(1-\delta^{\alpha}\right)}\right]^{\frac{\alpha\beta}{1-\alpha}}\left(1-\delta^{\alpha}\right)^{\frac{\alpha}{1-\alpha}}}_{=F_1}\cdot\underbrace{\left[1+\frac{\alpha\delta^{\alpha}\left(1-\phi_{O}\right)\left[\mu-\beta\left(1+\mu\right)\right]}{\gamma-\alpha\left(1-\beta\right)\mu+\phi_{O}\alpha\left[\mu-\beta\left(1+\mu\right)\right]}\right]}_{=F_2}.$$

As before, $\gamma,\mu\in(0,1)$ deliver Θ^{u} and $\gamma=\mu=1$ deliver Θ^{e}. The derivative of $\Theta^{l}(\beta)$ with respect to β is positive if

$$\Theta^{l\prime}(\beta)=\frac{\partial F_1}{\partial\beta}F_2+\frac{\partial F_2}{\partial\beta}F_1>0,$$

with

$$\begin{aligned}\frac{\partial F_1}{\partial\beta}=&\left(1-\delta^{\alpha}\right)^{\frac{\alpha}{1-\alpha}}\ln\left(1+\frac{\delta^{\alpha}}{\phi_{O}\left(1-\delta^{\alpha}\right)}\right)\frac{\alpha}{1-\alpha}\left[1+\frac{\delta^{\alpha}}{\phi_{O}\left(1-\delta^{\alpha}\right)}\right]^{\frac{\alpha\beta}{1-\alpha}}\\ \frac{\partial F_2}{\partial\beta}=&\frac{-\alpha\delta^{\alpha}\left(1-\phi_{O}\right)\left(1+\mu\right)\left[\gamma-\alpha\left(1-\beta\right)\mu+\phi_{O}\alpha\left[\mu-\beta\left(1+\mu\right)\right]\right]}{\left(\gamma-\alpha\left(1-\beta\right)\mu+\phi_{O}\alpha\left[\mu-\beta\left(1+\mu\right)\right]\right)^{2}}\\ &-\frac{\alpha\delta^{\alpha}\left(1-\phi_{O}\right)\left[\mu-\beta\left(1+\mu\right)\right]\left[\alpha\mu-\phi_{O}\alpha\left(1+\mu\right)\right]}{\left(\gamma-\alpha\left(1-\beta\right)\mu+\phi_{O}\alpha\left[\mu-\beta\left(1+\mu\right)\right]\right)^{2}}.\end{aligned}$$

$\Theta^{l\prime}(\beta)>0$ can be simplified to give

$$\ln\left(1+\frac{\delta^{\alpha}}{\phi_{O}\left(1-\delta^{\alpha}\right)}\right)\Omega\left(\beta,\mu,\gamma\right)>\left[\gamma\left(1+\mu\right)-\alpha\mu\right]\left(1-\alpha\right)\left(1-\phi_{O}\right)\delta^{\alpha}$$

where

$$\Omega\left(\beta,\mu,\gamma\right)=\underbrace{\left[\gamma-\alpha\mu\left(1-\phi_{V}\right)+\alpha\beta\left[\mu-\left(1+\mu\right)\phi_{V}\right]\right]}_{\tau_V}\underbrace{\left[\gamma-\alpha\mu\left(1-\phi_{O}\right)+\alpha\beta\left[\mu-\left(1+\mu\right)\phi_{O}\right]\right]}_{\tau_O}.$$

The strategy is now to show that Ω strictly decreases in β and then to plug in the minimum value $\Omega\left(\beta=1,\mu,\gamma\right)$ and show that the relationship still holds at this point. The two multiplicative terms τ_V and τ_O in Ω are symmetric except for the bargaining power parameter ϕ_k, so that

$$\frac{\partial\tau_k}{\partial\beta}=\alpha\left[\mu-\left(1+\mu\right)\phi_k\right]<0,\ k\ \in\ \{V,O\}.$$

To see this note that $\frac{\partial(\mu-(1+\mu)\phi_k)}{\partial\mu}=1-\phi_k>0$. The term therefore reaches its maximum at $\mu=1$, where it becomes $1-2\phi_k$, which is negative because $\phi_k>\frac{1}{2}$ by assumption. To determine the sign of $\frac{\partial\Omega}{\partial\beta}$, we need to determine the sign of τ_k, which can be rewritten as

$$\tau_k=\gamma-\alpha\left[\beta\phi_k+\left(1-\beta\right)\mu\left(1-\phi_k\right)\right].$$

The term in brackets can be shown to be smaller than ϕ_k because $\phi_k>\frac{1}{2}$. Therefore the assumption that $\gamma>\phi_V$ from the existence proof is sufficient to ensure a positive τ_k. Maintaining $\gamma>\phi_V$, it follows that

under both ethical and unethical production, τ_k is positive. This implies that

$$\frac{\partial \Omega\left(\beta,\mu,\gamma\right)}{\partial \beta} = \frac{\partial \tau_V}{\partial \beta}\tau_O + \frac{\partial \tau_O}{\partial \beta}\tau_V < 0.$$

It follows that Ω attains its smallest value within the admissible range of β at $\beta = 1$. Plugging in $\beta = 1$ into Ω eliminates μ from the function and yields

$$\Omega\left(\beta = 1,\gamma\right) = \left(\gamma - \alpha\phi_V\right)\left(\gamma - \alpha\phi_O\right).$$

Note that the assumption $\gamma > \phi_V$ ensures that both factors are positive because $\phi_O < \phi_V$. Expressing ϕ_V in terms of ϕ_O and inserting this for Ω in $\Theta^l(\beta)$ and rearranging then yields

$$\vartheta(\delta) = \ln\left(1 + \frac{\delta^\alpha}{\left(1-\delta^\alpha\right)\phi_O}\right) - \underbrace{\frac{\left[\gamma\left(1+\mu\right) - \alpha\mu\right]\left(1-\alpha\right)\left(1-\phi_O\right)\delta^\alpha}{\left[\gamma - \alpha\left(\phi_O + \delta^\alpha\left(1-\phi_O\right)\right)\right]\left(\gamma - \alpha\phi_O\right)}}_{=\Omega(\beta=1,\gamma)} \overset{!}{>} 0$$

To show that $\vartheta(\delta) > 0$ for all $\delta \in (0,1)$, note that $\vartheta(\delta = 0) = 0$ so that $\vartheta(\delta) > 0$ if $\vartheta'(\delta) > 0$. The first derivative of ϑ with respect to δ can be expressed as

$$\begin{aligned}\frac{\partial \vartheta}{\partial \delta} =& \frac{\alpha\delta^{\alpha-1}}{\left(1-\delta^\alpha\right)\left[\delta^\alpha + \phi_O\left(1-\delta^\alpha\right)\right]} - \frac{\left[\gamma\left(1+\mu\right) - \alpha\mu\right]\left(1-\alpha\right)\left(1-\phi_O\right)}{\left(\gamma - \alpha\phi_O\right)\left[\gamma - \alpha\left(\phi_O + \delta^\alpha\left(1-\phi_O\right)\right)\right]^2} \\ &\cdot\left(\alpha\delta^{\alpha-1}\right)\left\{\left[\gamma - \alpha\left(\phi_O + \delta^\alpha\left(1-\phi_O\right)\right)\right] + \alpha\delta^\alpha\left(1-\phi_O\right)\right\} \overset{!}{>} 0.\end{aligned}$$

This can be simplified further to give

$$\left(\gamma - \alpha\phi_V\right)^2 \overset{!}{>} \left[\gamma\left(1+\mu\right) - \mu\alpha\right]\left(1-\alpha\right)\left(1-\phi_V\right)\phi_V \equiv M(\mu).$$

Now note that $\frac{\partial M}{\partial \mu} = \left(\gamma - \alpha\right)\left(1-\alpha\right)\left(1-\phi_V\right)\phi_V$. The sign of the derivative depends on the relationship between γ and α.

Case 1 Consider case 1 where $\gamma < \alpha$ and so $\frac{\partial M}{\partial \mu} < 0$. This implies that for $\mu \in (0,1)$, $M(\mu)$ attains a maximum in the corner case of $\mu = 0$. For the inequality above to hold it is therefore sufficient to prove that

$$\left(\gamma - \alpha\phi_V\right)^2 > \gamma\left(1-\alpha\right)\left(1-\phi_V\right)\phi_V. \tag{A.7}$$

Simplifying and solving for α equivalently gives

$$\alpha^2\phi_V^2 - \alpha\gamma\phi_V\left(1+\phi_V\right) + \gamma\left[\gamma - \left(1-\phi_V\right)\phi_V\right] > 0.$$

The discriminant term of this quadratic equation is given by

$$\left(1+\phi_V\right)^2\phi_V^2\gamma^2 - 4\phi_V^2\gamma\left[\gamma - \phi_V\left(1-\phi_V\right)\right].$$

Simplification shows that the discriminant term is negative if $\gamma > \frac{4\phi_V}{3+\phi_V}$ so that (A.7) has no roots and is thus always positive. Because $\frac{4\phi_V}{3+\phi_V} > \phi_V \;\forall\; \phi_V \in (0,1)$, the inequality (A.7) holds for all $\alpha \in (0,1)$ when

$\gamma > \frac{4\phi_V}{3+\phi_V} > \phi_V$ and $\gamma < \alpha$.

We have previously imposed $\gamma > \phi_V$ to guarantee existence of β_u. Now consider values of γ between ϕ_V and $\frac{4\phi_V}{3+\phi_V}$. (A.7) has roots in this parameter range. For (A.7) to hold for all α for some $\gamma < \frac{4\phi_V}{3+\phi_V}$, we would need the smaller of the two roots of (A.7) to be larger than 1, which requires

$$\gamma\left(1+\phi_V\right) - 2\phi_V > \sqrt{\left(1+\phi_V\right)^2\gamma^2 - 4\gamma\left[\gamma - \left(1-\phi_V\right)\phi_V\right]}. \tag{A.8}$$

The right-hand side is the discriminant term and is positive because we consider values of $\gamma < \frac{4\phi_V}{3+\phi_V}$. The left-hand side is only positive if $\gamma > \frac{2\phi_V}{1+\phi_V}$, which is larger than $\frac{4\phi_V}{3+\phi_V}$. This implies that in the range of values of γ we consider here, the left-hand side is always negative and so (A.8) never holds for these values. In the rest of the proof, we must therefore impose the stricter condition $\gamma > \frac{4\phi_V}{3+\phi_V}$.

Case 2 Consider case 2 where $\gamma > \alpha$ and so $\frac{\partial M}{\partial \mu} > 0$. This implies that for $\mu \in (0,1)$, $M(\mu)$ attains a maximum at the corner case $\mu = 1$. The relationship to be shown now is

$$\left(\gamma - \alpha\phi_V\right)^2 - \left(2\gamma - \alpha\right)\left(1-\alpha\right)\left(1-\phi_V\right)\phi_V > 0. \tag{A.9}$$

Note first that for the left-hand side to be increasing in γ, it has to hold that $\gamma > \phi_V\left[1 - \phi_V\left(1-\alpha\right)\right]$. Because the term in brackets is smaller than 1, this is true for all $\gamma > \frac{4\phi_V}{3+\phi_V} \geq \phi_V$. It is therefore sufficient to show that (A.9) holds at the minimum level of γ. In this case we assume $\gamma > \alpha$ and impose $\gamma > \frac{4\phi_V}{3+\phi_V} > \phi_V$. Three sub-cases have to be covered.

Case 2a: $\gamma > \alpha > \frac{4\phi_V}{3+\phi_V}$ The minimum value γ can take here is α. Plugging in α for γ in (A.9) and simplifying gives that (A.9) holds when $\alpha > \phi_V$, which is true in this sub-case because $\frac{4\phi_V}{3+\phi_V} > \phi_V$.

Case 2b: $\gamma > \frac{4\phi_V}{3+\phi_V} > \alpha > \phi_V$ The minimum value γ can take here is $\frac{4\phi_V}{3+\phi_V}$. Case 2a has shown that if $\alpha > \phi_V$, (A.9) holds for $\gamma > \alpha$ which also holds in this case.

Case 2c: $\gamma > \frac{4\phi_V}{3+\phi_V} > \phi_V > \alpha$ Plugging in ϕ_V for γ in (A.9) results in the necessary condition of $\alpha < \phi_V$ for (A.9) to hold, which is true here. (A.9) therefore holds for $\gamma > \phi_V$ when $\phi_V > \alpha$. This includes $\frac{4\phi_V}{3+\phi_V} > \phi_V$. QED.

A.2.3 Relative size of the two integration cutoffs

We prove that $\beta_u > \beta_e$ by showing that (1) $\frac{\partial \beta_u}{\partial \mu} < 0$ for all $\mu \in (0,1]$ and $\gamma \in \left(\frac{4\phi_V}{3+\phi_V}, 1\right]$ and that (2) $\frac{\partial \beta_u}{\partial \gamma} < 0$ for all $\mu \in (0,1]$ and $\gamma \in \left(\frac{4\phi_V}{3+\phi_V}, 1\right]$. This includes the corner case of $\mu = \gamma = 1$, in which $\beta_u = \beta_e$. This implies that starting from the case $\beta_u = \beta_e$, any marginal decrease in either μ or γ increases β_u and continues to do so over the admissible range of the two parameters. We prove this using implicit differentiation of

$$\Theta^u(\beta_u) = F_1 \cdot F_2 = \left[1 + \frac{\delta^\alpha}{\phi_O\left(1-\delta^\alpha\right)}\right]^{\frac{\alpha\beta_u}{1-\alpha}} \left(1-\delta^\alpha\right)^{\frac{\alpha}{1-\alpha}} \cdot \left[1 + \frac{\alpha\delta^\alpha\left(1-\phi_O\right)\left[\mu - \beta_u\left(1+\mu\right)\right]}{\gamma - \alpha\left(1-\beta_u\right)\mu + \phi_O\alpha\left[\mu - \beta_u\left(1+\mu\right)\right]}\right] = 1$$

with respect to μ and γ.

Derivative of β_u with respect to μ First note that

$$\frac{\partial F_1}{\partial \mu} = (1-\delta^\alpha)^{\frac{\alpha}{1-\alpha}} \ln\left(1+\frac{\delta^\alpha}{\phi_O(1-\delta^\alpha)}\right)\left[1+\frac{\delta^\alpha}{\phi_O(1-\delta^\alpha)}\right]^{\frac{\alpha\beta_u}{1-\alpha}} \frac{\alpha}{1-\alpha}\frac{\partial \beta_u}{\partial \mu},$$

$$\begin{aligned}\frac{\partial F_2}{\partial \mu} = &\frac{\alpha\delta^\alpha(1-\phi_O)\left[1-\frac{\partial\beta_u}{\partial\mu}-\left(\beta_u+\mu\frac{\partial\beta_u}{\partial\mu}\right)\right]\left[\gamma-\alpha(1-\beta_u)\mu+\phi_O\alpha\left[\mu-\beta_u(1+\mu)\right]\right]}{\left\{\gamma-\alpha(1-\beta_u)\mu+\phi_O\alpha\left[\mu-\beta_u(1+\mu)\right]\right\}^2}\\ &-\frac{\alpha\delta^\alpha(1-\phi_O)\left[\mu-\beta_u(1+\mu)\right]\left[\alpha\left(\beta_u+\mu\frac{\partial\beta_u}{\partial\mu}\right)-\alpha+\phi_O\alpha\left[1-\frac{\partial\beta_u}{\partial\mu}-\left(\beta_u+\mu\frac{\partial\beta_u}{\partial\mu}\right)\right]\right]}{\left\{\gamma-\alpha(1-\beta_u)\mu+\phi_O\alpha\left[\mu-\beta_u(1+\mu)\right]\right\}^2}\end{aligned}$$

and that $\frac{\partial 1}{\partial \mu}=0$. Combining the terms to write $\frac{\partial F_1}{\partial \mu}F_2 + F_1\frac{\partial F_2}{\partial \mu}=0$ and simplification by multiplying through with the denominator term from $\frac{\partial F_2}{\partial \mu}$ gives that

$$\begin{aligned}&\ln\left(1+\frac{\delta^\alpha}{\phi_O(1-\delta^\alpha)}\right)\Omega(\beta_u,\gamma,\mu)\frac{\partial\beta_u}{\partial\mu}\\ =&\delta^\alpha(1-\phi_O)(1-\alpha)\left\{\left[\alpha\left(\beta_u+\mu\frac{\partial\beta_u}{\partial\mu}\right)-\alpha+\alpha\phi_O\left[1-\frac{\partial\beta_u}{\partial\mu}-\left(\beta_u+\mu\frac{\partial\beta_u}{\partial\mu}\right)\right]\right]\left[\mu-\beta_u(1+\mu)\right]\right.\\ &\left.-\left[1-\frac{\partial\beta_u}{\partial\mu}-\left(\beta_u+\mu\frac{\partial\beta_u}{\partial\mu}\right)\right]\left[\gamma-\alpha(1-\beta_u)\mu+\phi_O\alpha\left[\mu-\beta_u(1+\mu)\right]\right]\right\},\end{aligned}$$

where $\Omega(\beta_u,\gamma,\mu)$ is defined as above. The term in braces can then be simplified and the expression becomes

$$\begin{aligned}&\ln\left(1+\frac{\delta^\alpha}{\phi_O(1-\delta^\alpha)}\right)\Omega(\beta_u,\gamma,\mu)\frac{\partial\beta_u}{\partial\mu}\\ =&\delta^\alpha(1-\phi_O)(1-\alpha)\left\{\frac{\partial\beta_u}{\partial\mu}\left[\gamma(1+\mu)-\alpha\mu\right]-\gamma(1-\beta_u)-\alpha\beta_u^2(1-\mu)\right\},\end{aligned}$$

which can be rearranged to

$$\begin{aligned}&\frac{\partial\beta_u}{\partial\mu}\left[\delta^\alpha(1-\phi_O)(1-\alpha)\left[\gamma(1+\mu)-\alpha\mu\right]-\ln\left(1+\frac{\delta^\alpha}{\phi_O(1-\delta^\alpha)}\right)\Omega(\beta_u,\gamma,\mu)\right]\\ =&\delta^\alpha(1-\phi_O)(1-\alpha)\left[\gamma(1-\beta_u)+\alpha\beta_u^2(1-\mu)\right].\end{aligned}$$

Notice that the term on the right-hand side is positive for the admissible ranges of the parameters. In particular, it is also positive for $\gamma,\mu\in(0,1]$. To get $\frac{\partial\beta_u}{\partial\mu}<0$, we need that the term in square brackets on the left-hand side is negative, or equivalently that

$$\delta^\alpha(1-\phi_O)(1-\alpha)\left[\gamma(1+\mu)-\alpha\mu\right] < \ln\left(1+\frac{\delta^\alpha}{\phi_O(1-\delta^\alpha)}\right)\Omega(\beta_u,\gamma,\mu).$$

Because we assume $\phi_k > \frac{1}{2}$, we know from the uniqueness proof in Section A.2.2 that $\Omega(\beta_u, \gamma, \mu)$ has a minimum at $\beta_u = 1$. Plugging in $\beta_u = 1$ and rearranging shows that we need

$$\ln\left(1 + \frac{\delta^\alpha}{\phi_O(1-\delta^\alpha)}\right) - \frac{\delta^\alpha(1-\phi_O)(1-\alpha)[\gamma(1+\mu)-\alpha\mu]}{\Omega(\beta_u = 1, \gamma, \mu)} \equiv \vartheta(\delta) > 0$$

to obtain $\frac{\partial \beta_u}{\partial \mu} < 0$. In the uniqueness part in Section A.2.2 above it has been shown that the condition above holds if $\gamma > \frac{4\phi_V}{3+\phi_V}$, and in particular this holds when $\mu = \gamma = 1$. $\frac{\partial \beta_u}{\partial \mu} < 0$ for $\mu \in (0,1]$ and $\gamma \in \left(\frac{4\phi_V}{3+\phi_V}, 1\right]$ implies that β_u is increasing in the unethical cost advantage $1-\mu$ for any of these values of μ and γ.

Derivative of β_u with respect to γ First note that

$$\frac{\partial F_1}{\partial \gamma} = F_1 \ln\left(1 + \frac{\delta^\alpha}{\phi_O(1-\delta^\alpha)}\right)\frac{\alpha}{1-\alpha}\frac{\partial \beta_u}{\partial \gamma}$$

and

$$\frac{\partial F_2}{\partial \gamma} = \frac{-\alpha\delta^\alpha(1-\phi_O)[\mu-\beta_u(1+\mu)] - \alpha\delta^\alpha(1-\phi_O)[\gamma(1+\mu)-\alpha\mu]\frac{\partial \beta_u}{\partial \gamma}]}{(\gamma - \alpha(1-\beta_u)\mu + \phi_O\alpha[\mu - \beta_u(1+\mu)])^2}.$$

Combining those two derivatives in the equation $F_1\frac{\partial F_2}{\partial \gamma} + F_2\frac{\partial F_1}{\partial \gamma} = 0$ and solving for $\frac{\partial \beta_u}{\partial \gamma}$ gives

$$\frac{\partial \beta_u}{\partial \gamma} = \frac{(1-\alpha)\delta^\alpha(1-\phi_O)[\mu-\beta_u(1+\mu)]}{\ln\left(1+\frac{\delta^\alpha}{\phi_O(1-\delta^\alpha)}\right)\Omega(\beta_u,\mu,\gamma) - \delta^\alpha(1-\phi_O)(1-\alpha)[\gamma(1+\mu)-\alpha\mu]}.$$

The sign of the derivative is ambiguous. The denominator is positive, including the case of $\mu = \gamma = 1$, as can be seen from the uniqueness proof in Section A.2.2. The numerator is only negative if $\beta_u > \frac{\mu}{1+\mu}$, where $\frac{\mu}{1+\mu}$ reaches its maximum of $\frac{1}{2}$ at $\mu = 1$. Therefore, $\frac{\partial \beta_u}{\partial \gamma} < 0$ iff $\beta_u > \frac{1}{2}$.

The strategy is now to show that $\beta_e > \frac{1}{2}$. This will then imply that when $\mu = \gamma = 1$ and thus $\beta_u = \beta_e$, the numerator is negative and thus $\frac{\partial \beta_u}{\partial \gamma} < 0$. This then proves that starting from $\beta_u = \beta_e$, any decrease in γ increases β_u and does so for the whole range of admissible parameter values, i.e. $\mu \in (0,1]$ and $\gamma \in \left(\frac{4\phi_V}{3+\phi_V}, 1\right]$. To see this, consider the parameter condition needed to produce $\beta_e = \frac{1}{2}$ as the ethical integration cutoff. $\Theta^e\left(\beta_e = \frac{1}{2}\right) = 1$ after some algebra simplifies considerably to

$$\phi_O = D(\delta) \equiv \frac{1-\delta^\alpha}{2-\delta^\alpha}.$$

As $\delta, \alpha \in (0,1)$, $D(\delta)$ reaches its maximum of $\frac{1}{2}$ as $\delta \to 0$. This means that to have $\beta_e = \frac{1}{2}$, we need $\phi_O = D(\delta)$ with $D(\delta) < \frac{1}{2}$. This is ruled out by the initial assumption that $\phi_O > \frac{1}{2}$, which we carry over from Antràs (2003). We have now merely shown that $\beta_e = \frac{1}{2}$ is impossible under the imposed parameter restrictions. The proof is only complete if we show that any $\beta_e < \frac{1}{2}$ requires a value of ϕ_O whose maximum also lies below $\frac{1}{2}$. Therefore, we show that $\frac{\partial \beta_e}{\partial \phi_O} > 0$, implying that a decrease in β_e requires a reduction

in ϕ_O, c.p. Implicit differentiation yields that

$$\frac{\partial \beta_e}{\partial \phi_O} = \frac{\frac{\Omega(\beta_e)\beta_e\delta^\alpha}{\phi_O[\phi_O(1-\delta^\alpha)+\delta^\alpha]} + \delta^\alpha(1-\alpha)[1-2\beta_e][1-\alpha\beta_e]}{\ln\left(1+\frac{\delta^\alpha}{\phi_O(1-\delta^\alpha)}\right)\Omega(\beta_e) - \delta^\alpha(1-\phi_O)(1-\alpha)[2-\alpha]}.$$

The sign of the derivative is again determined by the sign of the numerator. The denominator is positive as has been shown in the uniqueness part of the proof in Section A.2.2. If $\beta_e \leq \frac{1}{2}$, the numerator is positive. Therefore, for $\beta_e \leq \frac{1}{2}$, a marginal decrease in β_e would require a decrease in ϕ_O. So in order to have a $\beta_e < \frac{1}{2}$ we require $\phi_O < D(\delta)$, which is ruled out by the initial assumption of $\phi_O > \frac{1}{2}$. QED.

A.3 Proof of Proposition 2

(i) For the proofs of Cases 2 and 3 note that the existence part of the proof of Proposition 1 specifies conditions for which $\Theta^u(\beta_u)$ and $\Theta^e(\beta_e)$ are *smaller* than 1 and *larger* than 0, respectively. Therefore, as long as these conditions hold, $\beta_e \in (0,1)$ and $\beta_u \in (0,1)$. Showing that $\beta_S \leq 0$ and $\beta_S = 1$ are possible within the admissible range of the parameters determining the cutoff proves the existence of Cases 2 and 3. With these preliminaries, it is unnecessary to consider partial derivatives of β_e and β_u with respect to μ and γ, because by Proposition 1, $\beta_e, \beta_u \in (0,1)$.

Case 3 For this case we show that as the unethical cost advantage goes to zero ($\mu \to 1$), $\beta_S \to -\infty$ so that unethical production is never chosen.

$$\lim_{\mu\to 1} \beta_S = \lim_{\mu\to 1}\left[1 - \frac{1-\gamma}{\alpha(1-\mu)}\right] = -\infty. \tag{A.10}$$

Case 2 For this case we show that $\beta_S \to 1$ as the threat of a consumer boycott goes to zero ($\gamma \to 1$) so that ethical production is never chosen.

$$\lim_{\gamma\to 1} \beta_S = \lim_{\gamma\to 1}\left[1 - \frac{1-\gamma}{\alpha(1-\mu)}\right] = 1 - \frac{0}{\alpha(1-\mu)} = 1. \tag{A.11}$$

Case 1 Consider some $\beta_S \in (-\infty, 1)$. Case 1 trivially exists if $\beta_e < \beta_S < \beta_u$. Case 1 also exists starting from any value of $\beta_S < \beta_e$ or $\beta_S > \beta_u$. If $\beta_S < \beta_e$, increasing $\gamma \to 1$ will necessarily move $\beta_S \to 1$, while $\beta_e, \beta_u \in (0,1)$. For some values of γ given μ and α, it must be the case that $\beta_e < \beta_S < \beta_u$. If $\beta_S > \beta_u$, increasing $\mu \to 1$ will necessarily move $\beta_S \to -\infty$, while $\beta_e, \beta_u \in (0,1)$. For some values of μ given γ and α, it must be the case that $\beta_e < \beta_S < \beta_u$.

(ii) Consider Case 1. Firms in sectors with headquarter intensity $\beta < \beta_e$ choose outsourcing and unethical production. In sectors with $\beta \in (\beta_e, \beta_S)$ unethical production and outsourcing continue to be chosen because β_u is the relevant cutoff for the headquarter in this case. For $\beta > \beta_S$, ethical production is chosen and with β_e as the relevant cutoff, also integration. This makes β_S the integration cutoff in this case so that $\bar{\beta} = \beta_S$.

Considering Case 2, firms in sectors with $\beta < \beta_u$ choose unethical production and outsourcing as β_u is the relevant cutoff for the headquarter in this case. For $\beta > \beta_u$, integration is chosen irrespective of the chosen technology, making β_u the relevant integration cutoff so that $\bar{\beta} = \beta_u$.

In Case 3, suppliers in sectors above β_S choose ethical production, so that there is integration above β_e and β_e is the integration cutoff so that $\bar{\beta} = \beta_e$.

(iii) This follows directly from (i) and (ii). QED.

A.4 Proof of Proposition 3

It has been shown in the proof of Proposition 1 in Section A.2.3 that $\frac{\partial \beta_u}{\partial \mu} < 0$ for $\mu \in (0, 1]$ and $\gamma \in \left(\frac{4\phi_V}{3+\phi_V}, 1\right]$. It follows directly that $\frac{\partial \beta_u}{\partial (1-\mu)} > 0$ for these parameter values. It has been shown in the proof of Lemma 1 in Section A.1 that $\frac{\partial \beta_S}{\partial (1-\mu)} > 0$. Moreover, it can be seen from equation (22) that β_e does not depend on μ or γ. Therefore, $\frac{\partial \beta_e}{\partial \mu} = 0$. QED.

A.5 Proof of Proposition 4

The proof follows closely the proof of Lemma 1 in terms of structure. When the headquarter can also set the technology of the match in addition to the organizational form, the key difference is that the headquarter takes the overall surplus into account when deciding between ethical and unethical production. The headquarter again compares the total cost savings from unethical production ΔC to the expected ethical revenue premium, which we now label $E[\Delta R]$, which is given by the sum of the suppliers and the headquarters revenue premium. Therefore, the term $(1 - \phi_k)$ on the right-hand side, which denoted the revenue share allocated to the supplier in the proof of Lemma 1 is now replaced by unity.

$$
\begin{aligned}
\Delta C &> E[\Delta R] \\
(c_m^e - c_m^u)\, m(\omega)_k^e &> (R(\omega)_k^e - E[R(\omega)_k^u]) \\
(c_m^e - c_m^u)\, m(\omega)_k^e &> (1-\gamma)\, R(\omega)_k^e \\
(c_m^e - c_m^u)\,(1-\beta)\, A\alpha^{\frac{1}{1-\alpha}} \frac{1-\phi_k}{c_m^e} \left[\left(\frac{c_h}{\phi_k}\right)^{\beta} \left(\frac{c_m^e}{1-\phi_k}\right)^{1-\beta}\right]^{-\frac{\alpha}{1-\alpha}} &> (1-\gamma)\, A\alpha^{\frac{\alpha}{1-\alpha}} \left[\left(\frac{c_h}{\phi_k}\right)^{\beta} \left(\frac{c_m^e}{1-\phi_k}\right)^{1-\beta}\right]^{-\frac{\alpha}{1-\alpha}} \\
\frac{c_m^e - c_m^u}{c_m^e}\,(1-\phi_k)\,(1-\beta)\,\alpha &> 1-\gamma
\end{aligned}
$$

Solving for β using the fact that $c_m^u = \mu c_m^e$ gives that when

$$
\beta < \beta_{H,k} = 1 - \frac{1-\gamma}{(1-\mu)\,\alpha\,(1-\phi_k)} < \beta_S, \tag{A.12}
$$

the headquarter will prefer unethical production. Note that this cutoff now depends on the organizational form of the firm. $\beta_{H,O} > \beta_{H,V}$ because $\phi_V > \phi_O$. Because $1 - \phi_k < 1$, both cutoffs are smaller than β_S from the baseline model.

Existence of the described pattern From Section A.3 we know that by letting $\gamma \to 1$, $\beta_S \to 1$ and by letting $\mu \to 1$, $\beta_S \to -\infty$. Because the new cutoffs $\beta_{H,k}$ with $k \in \{V, O\}$ differ from β_S only by a

positive factor in the denominator, the results from Proposition 2 can be directly applied to the cutoffs derived above. Therefore, there is a non-empty set of admissible values of γ, α and μ that ensures that for any $\phi_k > \frac{1}{2}$, there exists a range of $\beta \in (\beta_{H,O}, \beta_S) \wedge \beta \in (\beta_e, \beta_u)$. QED.

References

Acemoglu, D., P. Antràs, and E. Helpman (2007): "Contracts and Technology Adoption," *American Economic Review*, 97, 916–943.

Aichele, R. and G. Felbermayr (2015): "Kyoto and Carbon Leakage: An Empirical Analysis of the Carbon Content of Bilateral Trade," *Review of Economics and Statistics*, 97, 104–115.

Aldashev, G., M. Limardi, and T. Verdier (2015): "Watchdogs of the Invisible Hand: NGO Monitoring and Industry Equilibrium," *Journal of Development Economics*, 116, 28–42.

Aldashev, G. and T. Verdier (2009): "When NGOs Go Global: Competition on International Markets for Development Donations," *Journal of International Economics*, 79, 198–210.

Antràs, P. (2003): "Firms, Contracts, and Trade Structure," *The Quarterly Journal of Economics*, 118, 1375–1418.

Antràs, P. and D. Chor (2013): "Organizing the Global Value Chain," *Econometrica*, 81, 2127–2204.

Antràs, P. and E. Helpman (2004): "Global Sourcing," *Journal of Political Economy*, 112, 552–580.

——— (2008): "Contractual Frictions and Global Sourcing," in *The Organization of Firms in a Global Economy*, ed. by E. Helpman, D. Marin, and T. Verdier, Harvard University Press, 9–54.

Antràs, P. and S. R. Yeaple (2014): "Multinational Firms and the Structure of International Trade," in *Handbook of International Economics*, ed. by G. Gopinath, E. Helpman, and K. Rogoff, Elsevier, vol. 4, 55–130.

Antràs, P. (2016): *Global Production: Firms, Contracts, and Trade Structure*, Princeton University Press.

Baldwin, R. and J. Harrigan (2011): "Zeros, Quality, and Space: Trade Theory and Trade Evidence," *American Economic Journal: Microeconomics*, 3, 60–88.

Baron, D. P. (2001): "Private Politics, Corporate Social Responsibility, and Integrated Strategy," *Journal of Economics & Management Strategy*, 10, 7–45.

——— (2003): "Private Politics," *Journal of Economics & Management Strategy*, 12, 31–66.

——— (2010): "Morally-Motivated Self-Regulation," *American Economic Review*, 100, 1299–1329.

——— (2011): "Credence Attributes, Voluntary Organizations, and Social Pressure," *Journal of Public Economics*, 95, 1331–1338.

——— (2012): *Business and Its Environment*, Pearson, seventh ed.

——— (2016): "Self-Regulation and the Market for Activism," *Journal of Economics & Management Strategy*, 25, 584–607.

Baron, D. P. and D. Diermeier (2007): "Strategic Activism and Nonmarket Strategy," *Journal of Economics & Management Strategy*, 16, 599–634.

Basu, K. and Z. Tzannatos (2003): "The Global Child Labor Problem: What Do We Know and What Can We Do?" *World Bank Economic Review*, 17, 147–173.

Battaglini, M. and B. Harstad (2018): "The Political Economy of Weak Treaties," *Journal of Political Economy, forthcoming.*

Besley, T. and M. Ghatak (2007): "Retailing Public Goods: the Economics of Corporate Social Responsibility," *Journal of Public Economics*, 91, 1645–1663.

Bird, Y. Z., J. L. Short, and M. W. Toffel (2019): "Coupling Labor Codes of Conduct and Supplier Labor Practices: The Role of Internal Structural Conditions," *Organization Science*, 30, 847–867.

Bloomberg (2018): "Apple Supplier Workers Describe Noxious Hazards at China Factory," /https://www.bloomberg.com/news/articles/2018-01-16/workers-at-apple-supplier-catcher-describe-harsh-conditions, accessed September 19, 2019.

Broda, C. and D. E. Weinstein (2006): "Globalization and the Gains from Variety," *Quarterly Journal of Economics*, 121, 541–585.

Carluccio, J. and M. Bas (2015): "The Impact of Worker Bargaining Power on the Organization of Global Firms," *Journal of International Economics*, 96, 162–181.

China Labor Watch (2012): "Samsung Factory Exploiting Child Labor," /http://www.chinalaborwatch.org/upfile/2012_8_8/samsung8-27.pdf, accessed September 19, 2019.

——— (2015): "The Other Side of Fairy Tales. An Investigation of Labor Conditions at Five Chinese Toy Factories," /http://www.chinalaborwatch.org/upfile/2015_11_20/2015.11.20%20The%20Other%20Side%20of%20Fairy%20Tales%20-%20EN%20final.small.pdf, accessed September 19, 2019.

——— (2018): "Apple's Failed CSR Audit. A Report on Catcher Technology Polluting the Environment and Harming the Health of Workers," /http://www.chinalaborwatch.org/upfile/2018_01_12/20180116-1.pdf, accessed September 19, 2019.

Clean Clothes Campaign (2017): "Labour Without Liberty. Female Migrant Workers in Bangalore's Garment Industry," /https://cleanclothes.org/file-repository/resources-publications-labour-without-liberty-2013-female-migrant-workers-in-bangalores-garment-industry-full-version-1, accessed September 19, 2019.

Cone (2013): "2013 Cone Communications Social Impact Study," Research Report, Boston.

Copeland, B. R. and M. S. Taylor (1994): "North-South Trade and the Environment," *Quarterly Journal of Economics*, 109, 755–787.

——— (2004): "Trade, Growth, and the Environment," *Journal of Economic Literature*, 42, 7–71.

Corcos, G., D. M. Irac, G. Mion, and T. Verdier (2013): "The Determinants of Intrafirm Trade: Evidence from French Firms," *The Review of Economics and Statistics*, 95, 825–838.

Costinot, A., L. Oldenski, and J. Rauch (2011): "Adaptation and the Boundary of Multinational Firms," *Review of Economics and Statistics*, 93, 298–308.

Couttenier, M. and S. Hatte (2016): "Mass Media Effect on Non-Governmental Organizations," *Journal of Development Economics*, 123, 57–72.

Defever, F. and F. Toubal (2013): "Productivity, Relationship-Specific Inputs and the Sourcing Modes of Multinationals," *Journal of Economic Behavior and Organization*, 94, 245–357.

Distelhorst, G., J. Hainmueller, and R. M. Locke (2017): "Does Lean Improve Labor Standards? Management and Social Performance in the Nike Supply Chain," *Management Science*, 63, 707–728.

Distelhorst, G. and R. M. Locke (2018): "Does Compliance Pay? Social Standards and Firm-level Trade," *American Journal of Political Science*, 62, 695–711.

Eckel, C. and H. Egger (2009): "Wage Bargaining and Multinational Firms," *Journal of International Economics*, 77, 206–214.

Egels-Zandén, N. (2007): "Suppliers' Compliance with MNCs' Codes of Conduct: Behind the Scenes at Chinese Toy Suppliers," *Journal of Business Ethics*, 75, 45–62.

Egorov, G. and B. Harstad (2017): "Private Politics and Public Regulation," *Review of Economic Studies*, 84, 1652–1682.

Feddersen, T. J. and T. W. Gilligan (2001): "Saints and Markets: Activists and the Supply of Credence Goods," *Journal of Economics and Management Strategy*, 10, 149–171.

Flammer, C. (2013): "Corporate Social Responsibility and Shareholder Reaction: The Environmental Awareness of Investors," *Academy of Management Journal*, 56, 758–781.

Fortune (2015): "Hasbro, Mattel Toy Suppliers Slammed in Labor Report," `/http://fortune.com/2015/11/22/hasbro-mattel-china-labor/`, accessed September 19, 2019.

Fu, Q., J. Gong, and I. Png (2018): "Law, Social Responsibility, and Outsourcing," *International Journal of Industrial Organization*, 57, 114–146.

Greenpeace (2011): "Dirty Laundry. Unravelling the Corporate Connections to Toxic Water Pollution in China," `/http://www.greenpeace.org/eastasia/Global/international/publications/toxics/Water%202011/dirty-laundry-report.pdf`, accessed September 19, 2019.

——— (2012): "Toxic Threads: Putting Pollution on Parade. How Textile Manufacturers are Hiding Their Toxic Trail," `/http://www.greenpeace.org/eastasia/Global/international/publications/toxics/Water%202012/ToxicThreads02.pdf`, accessed September 19, 2019.

——— (2016): "The Detox Catwalk 2016. Who's On the Path to Toxic-Free Fashion?" `/https://www.greenpeace.org/archive-international/en/campaigns/detox/fashion/detox-catwalk/`, accessed September 19, 2019.

Grossman, G. M. and E. Helpman (1994): "Protection for Sale," *American Economic Review*, 84, 833–850.

——— (2002): "Integration versus Outsourcing in Industry Equilibrium," *Quarterly Journal of Economics*, 117, 85–120.

Grossman, S. J. and O. D. Hart (1986): "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94, 691–719.

Guardian (2014): "Sustainable Sugar: Coca-Cola and BP Signed Up But Will It Go Mainstream?" `/https://www.theguardian.com/sustainable-business/2014/sep/15/sustainable-sugar-can-coca-cola-bp-shell-bonsucro`, accessed September 19, 2019.

——— (2017): "Pepsico, Unilever and Nestlé Accused of Complicity in Illegal Rainforest Destruction," /https://www.theguardian.com/environment/2017/jul/21/pepsico-unilever-and-nestle-accused-of-complicity-in-illegal-rainforest-destruction, accessed September 19, 2019.

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.

Hallak, J. C. (2006): "Product Quality and the Direction of Trade," *Journal of International Economics*, 68, 238–265.

Harrison, A. and J. Scorse (2010): "Multinationals and Anti-Sweatshop Activism," *American Economic Review*, 100, 247–273.

Hart, O. and J. Moore (1990): "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 98, 1119–1158.

Hatte, S. and P. Koenig (2018): "The Geography of NGO Activism against Multinational Corporations," *World Bank Economic Review, forthcoming.*

Hiscox, M. J. and N. Smyth (2011): "Is There Consumer Demand for Fair Labor Standards? Evidence from a Field Experiment," *SSRN Working Paper*, 1820642.

Human Rights Watch (2004): "Turning a Blind Eye. Hazardous Child Labor in El Salvador's Sugarcane Cultivation," /https://www.hrw.org/sites/default/files/reports/elsalvador0604full_1.pdf, accessed September 19, 2019.

Hummels, D. and P. J. Klenow (2005): "The Variety and Quality of a Nation's Exports," *American Economic Review*, 95, 704–723.

Innes, R. (2006): "A Theory of Consumer Boykotts under Symmetric Information and Imperfect Competition," *Economic Journal*, 116, 355–381.

KantarMillwardBrown (2017): "Brandz Top 100 Most Valuable Global Brands 2017," /http://brandz.com/admin/uploads/files/BZ_Global_2017_Report.pdf, accessed September 19, 2019.

King, B. G. and S. A. Soule (2007): "Social Movements as Extra-institutional Entrepreneurs: The Effect of Protests on Stock Price Returns," *Administrative Science Quarterly*, 52, 413–442.

Kitzmuller, M. (2012): "Is Better Information Always Good News? International Corporate Strategy and Regulation," *World Bank Policy Research Working Paper 6251.*

Kitzmuller, M. and J. Shimshack (2012): "Economic Perspectives on Corporate Social Responsibility," *Journal of Economic Literature*, 50, 51–84.

Kohler, W. and M. Smolka (2014): "Global Sourcing and Firm Selection," *Economics Letters*, 124, 411–415.

Krautheim, S. and T. Verdier (2016): "Offshoring with Endogenous NGO Activism," *Journal of International Economics*, 101, 22–41.

Kugler, M. and E. Verhoogen (2012): "Prices, Plant Size, and Product Quality," *Review of Economic Studies*, 79, 307–339.

Locke, R. M., T. Kochan, M. Romis, and F. Qin (2007a): "Beyond Corporate Codes of Conduct: Work Organization and Labour Standards at Nike's Suppliers," *Industrial Labour Review*, 146, 21–40.

Locke, R. M., F. Qin, and A. Brause (2007b): "Does Monitoring Improve Labor Standards? Lessons from Nike," *Industrial Labor Relations Review*, 61, 3–31.

Loureiro, M. and J. Lotade (2005): "Do Fair Trade and Eco-Labels in Coffee Wake Up the Consumer Conscience?" *Ecological Economics*, 53, 129–138.

Lyon, T. P. and S. Salant (2013): "Linking Public and Private Politics: Activist Strategy for Industry Transformation," Mimeo.

Melitz, M. J. (2003): "The Impact of Trade On Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71, 1695–1725.

New York Times (1998): "Nike Pledges to End Child Labor And Apply U.S. Rules Abroad," /https://www.nytimes.com/1998/05/13/business/international-business-nike-pledges-to-end-child-labor-and-apply-us-rules-abroad.html, accessed September 19, 2019.

——— (2017): "Detained Over Ivanka Trump Factory Inspection, China Labor Activist Speaks," /https://www.nytimes.com/2017/07/10/business/china-labor-ivanka-trump-shoes.html, accessed September 19, 2019.

Nike (1998): "Annual Report," /http://s1.q4cdn.com/806093406/files/doc_financials/1998/main_ar.html.

Nunn, N. and D. Trefler (2013): "Incomplete Contracts and the Boundaries of the Multinational Firm," *Journal of Economic Behavior and Organization*, 94, 330–344.

O'Rourke, D. (2005): "Market Movements: Nongovernmental Strategies to Influence Global Production and Consumption," *Journal of Industrial Ecology*, 9, 115–128.

Ouellet, M., J. L. Short, and M. W. Toffel (2015): "Codes in Context: How States, Markets, and Civil Society Shape Adherence to Global Labor Standards," *Regulation and Governance*, 9, 205–223.

Rainforest Action Network (2017): "Leuser Watch - Destroying Elephant Habitat While Breaching the Indonesian Government Moratorium on Forest Clearance for Palm Oil," /https://d3n8a8pro7vhmx.cloudfront.net/rainforestactionnetwork/pages/17392/attachments/original/1486495650/RAN_Leuser_Watch_PT_Agra_Bumi_Niaga_20170207.pdf, accessed September 19, 2019.

Ruwanpura, K. N. and N. Wrigley (2011): "The Costs of Compliance? View of Sri Lankan Apparel Manufacturers in Times of Global Economic Crisis," *Journal of Economic Geography*, 11, 1031–1049.

Short, J. L. (2013): "Self-Regulation in the Regulatory Void: "Blue Moon" or "Bad Moon"?" *Annals of the American Academy of Political and Social Science*, 649, 22–34.

Spence, A. M. (1973): "Job Market Signaling," *Quarterly Journal of Economics*, 87, 355–374.

——— (1974): *Market Signaling*, Harvard University Press.

Williamson, O. E. (1985): *The Economic Institutions of Capitalism*, Free Press.