

# Responsible Sourcing: The First Step Is the Hardest

Pia Ramchandani, Hamsa Bastani, Ken Moon

The Wharton School, University of Pennsylvania, Philadelphia, PA 19104  
{piar2, hamsab, kenmoon}@wharton.upenn.edu

Responsible sourcing is a priority for companies and consumers concerned with corporate social responsibility (CSR) in global supply chains. Most brands' product lines contain just a few products certified by third parties—which suggests that brands limit their efforts at ensuring that suppliers behave responsibly. In this paper, we examine a previously under-appreciated role of certifications: that certifications enable brands to learn about *how* to source responsibly. By successfully certifying even a single product, the certifying brand may enjoy positive, knowledge-based spillovers encouraging responsible sourcing throughout its product line. Using data on the responsible sourcing decisions of coffee brands in the \$48B US consumer market, we find that certifying brands' rates of CSR violations (adjusted for disparities in production volume and detection) are similarly low regardless of whether the brand's portfolio is 3% certified or 100% certified—consistent with learning-based spillover effects. Certifying brands' violation rates are an estimated 61-78% lower than for comparable brands that make no CSR claims. While we find that brands making their own uncertified, on-packaging CSR claims also exhibit low CSR violation rates, their low violation rates are nearly entirely explained by the countries from which they source. In contrast, certifying brands appear uniquely able to source responsibly even from within “high-risk” countries. Our work novelly suggests that prevalent dual-sourcing may surprisingly amplify, rather than limit, responsible sourcing in supply chains, and that certified sourcing valuably develops the pool of responsible suppliers in high-risk countries.

*Key words:* certifications, corporate social responsibility, human rights, labor abuse, sustainable sourcing

---

## 1. Introduction

Over recent decades, consumers, investors, governments, and stakeholders have intensified scrutiny on the social and environmental impacts of firms' business operations. As firms increasingly source from supply chains extending across international borders, significant concerns about the abuses taking place within the supply chains have followed suit. According to the International Labor Organization, global supply chains' 450 million workers include 168 million subject to child labor, 21 million victims of forced labor, and 12 million annual deaths from unsafe work environments (ILO 2018). In environmental impact, consumer good supply chains contribute to over 80% of water loss, 60% of greenhouse gas emissions, and 67% of deforestation (Consortium 2016).

In response to supply chain abuses, brands have created corporate social responsibility (CSR) initiatives, *i.e.*, voluntary undertakings to curb or mitigate the harmful impact of their activities (Lee and Tang 2018, Li and Wu 2017), that focus on responsible sourcing. To source responsibly,

the firms behind the brands commit to working with suppliers aligned with their stated CSR standards and goals, which can include minimizing child labor, improving working conditions, and decreasing environmental impacts such as deforestation.

However, actually sourcing responsibly is operationally challenging, requiring that firms enforce well functioning governance, visibility, and transparency throughout their global supply chains (Locke et al. 2007a, Pagell and Wu 2009, Sodhi and Tang 2018). Firms must develop ethical standards for supplier behavior, disseminate them to diverse suppliers, and then audit supplier practices, either internally or by engaging third-party certification agencies such as Fair Trade (Thorlakson et al. 2018). Finding or developing reliably responsible suppliers requires overcoming considerable geographical distances and cultural and organizational divides, making local knowledge valuable (Boström 2015). The lack of generally agreed-upon ethical standards further impedes the coordination of sourcing and audits (Nelson et al. 2018). Even then, suppliers may illegally subcontract (Caro et al. 2018) or deliberately evade ethical responsibilities by sourcing from regions with looser regulations (Abbasi 2017, Hasle and Jensen 2012). Finally, to generate market demand, firms must persuasively communicate their sourcing to socially responsible consumers through marketing and on-package labeling (Kraft et al. 2019, Buell et al. 2019).

On the other hand, once firms develop the requisite operational and institutional knowledge, they may enjoy diminished marginal costs when further expanding the scope of their responsible sourcing to additional products. This paper examines whether brands that certifying just a few products as part of their product lines accrue learning-based spillovers that enable them to also reduce violations in the non-certified sections of their supply chains. This question is especially relevant because dual-sourcing (*i.e.*, when a firm sources from both certified and non-certified channels) is a common sourcing strategy (Thorlakson et al. 2018); Nielsen RetailScanner data show that 21% of grocery brands dual-source, while only 5% of grocery sales are actually certified. If only third-party certified goods are reliably responsible—*i.e.*, if spillovers and reputational safeguards are absent or ineffectual—dual-sourcing and low certification levels pose clear concerns.

We address this empirical question using data on coffee brands' responsible sourcing decisions and labor and environmental CSR violations in the \$48B US consumer market. As our key result, we find that brands certifying the first few products appear to accrue spillover-inducing gains in operational knowledge about *how* to source responsibly. Specifically, when a coffee brand certifies even a small part of its product portfolio, namely above 3% in sales, it accrues positive spillovers enabling ethical sourcing throughout the uncertified rest of its portfolio. Surprisingly, the overall rates of CSR violations (adjusted for disparities in production activity and violation detection) are similarly low for certifying brands regardless of whether the brand's portfolio is 5% certified or 100% certified. Learning-based spillover effects from certifying are consistent with brands achieving

uniformly low violations over such widely varied but positive levels of product line certification—the benefit of learning arrives disproportionately in the first step of certifying a first product. In magnitude, we estimate that certifying brands' violation rates are 61-78% lower than for comparable brands lacking any CSR claims on their products. This result is robust to different specifications and matching on a variety of brand-specific controls.

Reflecting an additional benefit of the operational knowledge they obtain by certifying, certifying brands appear uniquely capable of sourcing responsibly from within "high-risk" countries. To be clear, we do not find that third-party certification is a necessary condition for sourcing responsibly: brands making their own uncertified, on-packaging CSR claims also exhibit low CSR violation rates. On the other hand, these brands' low violation rates are nearly entirely explained by the countries from which they source. In contrast, dual-sourcing companies with more than 3% in certified sales maintain low violations in both high risk and safe countries.

Taken together, our work makes several contributions. First, whereas an extensive literature raises concerns regarding brands' prevalent dual-sourcing, we examine an under-appreciated benefit from partially certifying a brand's product line. By enabling brands to learn the operational practices behind sourcing responsibly, even limited certification efforts generate spillover effects that allow firms to reduce CSR violations in their uncertified sourcing channels. Thus, we novelly argue that dual-sourcing may surprisingly amplify, rather than limit, responsible sourcing in supply chains. Second, the literature remains divided on whether brand reputations adequately ensure the credibility of firms' product labeling, especially when firms make uncertified CSR claims on their product packaging (Giovannucci et al. 2008, Darnall and Sides 2008). Our results indicate that reputations-based claims do appear to deliver when CSR-friendly sourcing options are readily available to brands; however, our evidence also supports that it is only the certified brands that undertake the additional challenge of sourcing responsibly from high-risk countries. It remains unclear whether market reputation alone could motivate brands to seek and develop ethical suppliers in high-risk regions. Lastly, our findings inform the role of ethical certification programs. To maximize impact on CSR outcomes, certification agencies should leverage spillovers by emphasizing the on-boarding of new firms through first-time certifications and developing the responsible supplier pool in high-risk countries.

### 1.1. Related Literature

Our work contributes to the literature on the role of third-party certifications in CSR. Enforcing CSR through certifications remains a challenge outside of the few markets where consumers *directly* benefit from supplier responsibility—*e.g.*, certifications for food safety and energy-efficient appliances (Waldman and Kerr 2014). In contrast, when certified products appeal only to *socially*

responsible consumers (*e.g.*, ethical fashion apparel and sustainably caught seafood), theory suggests that dual-sourcing can detriment CSR outcomes by segmenting the consumer market (Guo et al. 2015). Companies may dual-source to superficially appease consumers, who are often no more willing to pay for heavy investments in ethical production than for small ones (Trudel and Cotte 2009). Empirically, dual-sourcing is troublingly prevalent (Thorlakson et al. 2018).

An extensive literature examines the critical role of transparency and audits in achieving supplier compliance with CSR, where supply chain transparency (*i.e.*, a firm disclosing information to consumers, investors, and other stakeholders about its CSR compliance) goes beyond visibility (*i.e.*, the firm's monitoring of activities and compliance in its supply chain) (Sodhi and Tang 2018, Kraft et al. 2018). Joint audits and supplier list disclosures can beneficially penalize suppliers for their non-compliant behaviors (Kalkanci and Plambeck 2018b, Chen et al. 2019, Fang and Cho 2019) or induce appropriate monitoring by NGOs (Chen et al. 2018). Kim (2015) studies how regulators can increase transparency by encouraging firms to voluntarily disclose violations. However, there are pitfalls. Disclosure mandates can deter firms from learning about supplier violations (Kalkanci and Plambeck 2018a). Overly frequent audits may cause suppliers to direct their efforts towards evading audits rather than actually preventing harm in their operations (Plambeck and Taylor 2015), and suppliers under pressure may engage in unauthorized subcontracting (Caro et al. 2018). Lastly, CSR outcomes can be impeded when it is relatively difficult for brands to be transparent, *i.e.*, to credibly signal responsibility (Guo et al. 2017).

Moreover, a growing body of research suggests that closely working with suppliers to address managerial challenges and standards leads to improved CSR compliance. Porteous et al. (2015) find that reducing supplier violations is strongly associated with incentivizing supplier compliance through increased business, public recognition, and training. Corbett et al. (2005) and Castka et al. (2015) demonstrate long-lasting improvements in financial performance from the adoption of effective quality management and environmental standards. In collaboration with Nike, Locke et al. (2007b) illustrate that monitoring suppliers alone generally fails to improve poor factory working conditions, whereas assistance in identifying and addressing operational root causes (*e.g.*, scheduling) can improve conditions. Amengual and Distelhorst (2019) and Boudreau (2020) corroborate the role of management practices in improving labor and safety compliance in garment factories, and Bloom et al. (2010) demonstrate that better managed firms have significantly lower greenhouse gas emissions. Distelhorst et al. (2015) and Toffel et al. (2015) emphasize the role of the surrounding social and regulatory institutions in creating compliance, while Bastani and de Zegher (2019) find that banning supplier practices related to harmful behavior may reduce violations. Learning-by-doing also poses an important hurdle for suppliers in adopting sustainable practices (Akkaya et al. 2020).

Finally, contractual arrangements and market institutions can facilitate CSR. By incentivizing community monitoring and enforcement, de Zegher et al. (2018) propose contracts to lessen deforestation in palm oil supply chains. Levi et al. (2019) connect agricultural wholesale marketplaces onto a common platform to improve small farmer welfare. Babich and Tang (2012) study inspection and market response as two alternative forms of contractual discipline, and Chen and Lee (2016) and Cho et al. (2019) design optimal contracts that combine supplier screening mechanisms and contingency payments. Additional research on contracts incorporates dynamics and multiple tiers (Huang et al. 2017, Lewis et al. 2019, Zhang et al. 2019). Orsdemir et al. (2019) focuses on supply chain structure and integration.

The remainder of the paper proceeds as follows. Section 2 describes how we sourced and matched our dataset, as well as its limitations. Section 3 discusses sourcing spillovers, suggested both by data and first-hand sources. Section 4 covers our empirical methodology and results, and Section 5 expands on implications.

## 2. Dataset

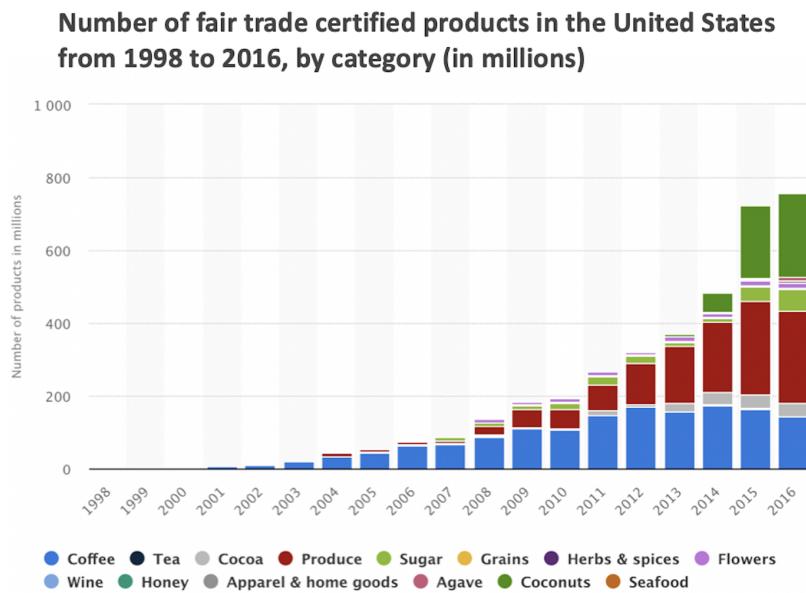
Our data covers sourcing decisions and metrics of supply chain responsibility in the \$48B US coffee market. The data integrate nearly 1TB of data from several sources.

*Choice of industry.* We choose to study the coffee industry for three reasons. First, the coffee market is highly mature in terms of brands' sourcing behavior. Coffee has a well established history of certifications, as shown in Figure 1 for the number of coffee products certified by Fair Trade. Sourcing strategies have been relatively stable for years, *e.g.*, we see very few changes in certification uptake during our study period. In addition, coffee has a significant representation of companies making “self-made” claims around ethical sourcing (see Figure 2(b)). This allows us to examine whether products bearing reputation-based, self-made claims are in fact responsibly sourced in comparison to certified products. This is important because self-made claims are becoming an increasingly prominent strategy with 31% of global grocery sales including self-made on-pack ethical claims in 2016.<sup>1</sup>

Second, the problem of responsibly sourcing coffee production is relevant in practice, because the industry is marred by significant ethical lapses. Even today, many coffee plantations fail to abide by labor laws, provide poor and hazardous working conditions, and exploit migrant and child laborers (Zamora 2013, US Department of Labor 2013). Regarding coffee’s environmental impact, millions of acres of forest have been cleared through illegal deforestation, destroying natural habitats (WWF

<sup>1</sup> We performed preliminary data collection in the chocolate industry (see Appendix E), but could not carry out a similar analysis as we did for coffee due to the limited penetration of dual-sourcing strategies; we do, however, find evidence suggesting similarly strong learning-based spillovers, indicating that insights from the coffee industry may generalize to other industries as they mature.

2007, The Guardian 2011). Finally, coffee products delineate into relatively homogenous specialty and commodity product segments, which allows us to more effectively control for product differences. Certified goods virtually always fall into the specialty segment. Because coffee certifications address only the subset of consumers who are socially responsible, dual-sourcing is natural and indeed prevalent in the coffee market, which is of interest for our study.



**Figure 1 Fair Trade certified products by category and year**



**Figure 2 Coffee Product Packaging with Sourcing-related Labels**

*Data sources.* Data regarding the studied brands' environmental, social and corporate governance (ESG) violations are sourced from RepRisk, which is a data provider that daily aggregates and cross-checks negative ESG incidents identified by over 90,000 news, media, regulatory, and commercial information sources published in over 20 languages. The violations data coverage is

comprehensive, recording and classifying all detected violation events for more than 140,000 companies over the span of 2006 to 2016. Data regarding the brands' sourcing strategies are sourced from Nielsen Label Insights, as derived from consumer product packaging and labels in 2016. For each individual product UPC, the data's label attributes granularly classify the on-packaging claims and certification endorsements. By identifying individual coffee products' claims and certifications, we are able to reconstruct and study brands' sourcing decisions at the level of product portfolios. As our last data of direct interest, we also obtain data on the brands' first-tier, cross-border suppliers from the Panjiva platform. These suppliers are the final international stop in the brands' coffee supply chains before the sourced coffee enters the US. Panjiva extracts data from import documents (*e.g.*, bills of lading) covering over one billion transactions by over 8 million companies across 195 countries.

Two additional pieces of data are used to appropriately normalize brands' detected violation counts within our analysis. First, the Nielsen Retail Scanner dataset supplies product-level retail sales data, using which we derive brands' coffee-based revenues and unit sales.<sup>2</sup> Second, we use data from LexisNexis, one of the world's largest electronic databases for legal and journalistic documents, to measure the volume of general news related to each company during each year from 2006 to 2016. As described below, we use these data to adjust brands' observed violation counts to account for their volumes of product sold, their speciality-versus-commodity mixes as captured by price points, and the levels of news scrutiny they receive.

The final dataset contains 413 observations where each observation represents one year of data for one particular coffee company during the time period 2006-2016. Some companies were launched during this timeframe, and therefore only contain data for a subset of the years. Each observation contains information on the number of ESG violations, country-based sourcing, scale of production, and news visibility for a particular company in that year. The final combined dataset covers 59% of the coffee sales in the US consumer market.

## 2.1. Outcome Variable

**2.1.1. ESG violation events.** Like Li and Wu (2017), our primary outcome variable is the supply chain ESG violation counts for the company-year collected by RepRisk. RepRisk screens over 500,000 documents every day across 90,000 public sources and stakeholders in 20 languages to identify violations. Public sources include print media, online media, social media, government bodies, regulators, think tanks, newsletters, and other online sources. These sources include coverage at the local, national, regional, and international levels. RepRisk uses a combination of

<sup>2</sup> The Retail Scanner dataset includes weekly volume, pricing and store conditions generated by point-of-sale systems from approximately 35,000 participating grocery, drug, mass merchandiser and other stores from 2006-2016. We use the Open Grocery Store Database to map product UPC codes and brands from the Retail Scanner data to their respective parent companies.

machine learning and manual review to tag incidents in 28 violation categories.<sup>3</sup> Since we focus on responsible sourcing, we filtered events by the available ESG violation categories to only include those related to sourcing, *e.g.*, we do not consider events solely relating to tax evasion or executive compensation. Table 1 lists the violation categories used in our study.

We later control for differences in violation detection rates based on company news visibility and size. Because different companies source from different locales, we additionally recognize that the definitions of violations can effectively differ across jurisdictions. Thus, certain activities may not be deemed violations simply because they occur in loosely regulated countries. First, this classification issue is addressed when companies are uniformly held to international or norm-based standards for ESG issues. Such standards are increasingly common and supported by the monitoring of NGOs and the media. Second, it is expected that irresponsible supplier behavior is more prevalent in loosely regulated nations. Because loose regulations relatively undercount violations, we may be concerned that the differences between countries may be washed out. However, when we rank our data sample's countries by violation rates, the order reassuringly accords with the views of expert observers. Lastly, we carry out regressions that control for companies' country-based sourcing. By focusing on within-country variation, these regressions validate whether certifying companies source more responsibly than their peers sourcing from the same countries.

We therefore analyze 2,216 eligible ESG events occurring between 2006-2016 for the companies in our study. For each company-year data observation, we aggregate the brand's detected violations to evaluate its overall ethical production performance.

Category	Issue categories	% of Events Tagged
<b>Environment</b>	climate change, GHG emissions, global pollution, impacts on landscapes, ecosystems, and/or biodiversity, local pollution, overuse and wasting of resources, waste issues	62%
<b>Social</b>	child labor, forced labor, human rights abuses and corporate complicity, impacts on communities, occupational health and safety issues, poor employment conditions	77%
<b>Governance</b>	violation of international standards, discrimination in employment, freedom of association and collective bargaining, social discrimination	17%

**Table 1 ESG issue categories that qualify brands' violation events for inclusion in the study**

Each studied violation event falls into one or more ESG issue categories listed above. An event can be tagged as a violation in multiple categories, so the % column adds up to more than 100%.

<sup>3</sup> Once an incident is flagged using natural language processing, analysts manually investigate to verify findings, remove duplicates, and classify the events into ESG categories. Incidents are also assigned proprietary scores to quantify the severity and reach of the event.

## 2.2. Study Variables

**2.2.1. Sourcing strategies.** We examine brands' sourcing strategies at the product portfolio level. We classify brands into three types of sourcing strategies, which are defined based on the sourcing composition of a brand's product line:

1. **Certified or dual-sourcing brands** certify at least 3% of the products in their product portfolios, through the endorsement of a major third-party certifier for ESG-related sourcing.<sup>4</sup>

2. **Brands relying on self-made claims** do not fall into category 1 and label at least one product with certification or the brands' own on-packaging claims concerning ESG-related sourcing (the claim therefore lacks external auditing or accreditation). For these brands to source responsibly, they must develop their own internal sourcing standards and programs, although they may outsource audits or other aspects of implementation.

3. **Brands with no responsible sourcing claims** lack ESG-related labels on any of their products.

This classification scheme allows us to differentiate between the effects of certification experience and of independent reputational branding. At various points in the study, we more closely consider the proportion of certified products within dual-sourcing brands' portfolios. In particular, we study the effects of certifying smaller (or larger) fractions of dual-sourcing brands' product lines.

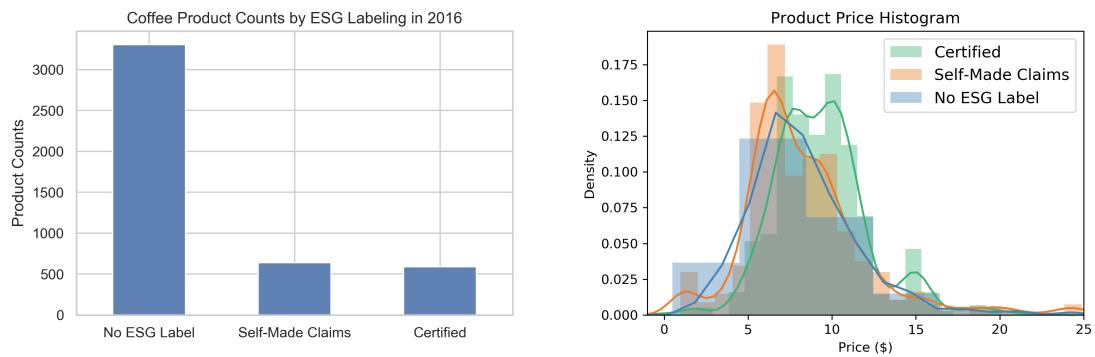
We classify the data's 4,534 unique coffee products into “certified”, “self-made claims”, and “no ESG labeling” as follows. From Nielsen LabelInsights data on product UPC labels' claims and third-party endorsements (*e.g.*, Fair Trade or Rainforest Alliance symbol), we filter down to labels relating to ethical sourcing. To keep our focus on sustainability, we omit certifications, such as the Organic standard, that solely concern growing practices. We identify 12 unique formal certification program labels (see list in Appendix A) and 176 self-made ESG claim labels. The Nielsen LabelInsights data on product labels were produced for 2016 only, therefore we identify brands' sourcing strategies based on their products' labeling in 2016. Because 70% of dual-sourcing companies were certified prior to 2006, we do not expect significant changes to their sourcing strategies during our study period (alleviating concerns about products that have adopted certification but not yet portrayed this fact through product packaging).

If a product bears a major certifier's third-party endorsement label, we classify it as a certified product.<sup>5</sup> If a brand labels a product with an uncertified ESG sourcing claim, the product is classified under “self-made claims”. All other products are designated as bearing no ESG labeling. After classifying each product, we classify brands into the three sourcing categories described above.

<sup>4</sup> The motivation for considering a 3% cutoff is presented in Section 3.

<sup>5</sup> We assume that third-party certified products display their certifications on on-package labeling to burnish the brand's image among socially responsible consumers (*see, e.g.*, Anagnostou et al. 2015).

Of the 4,534 products, 13% are certified, 14% have self-made claims, while the majority have no ESG labeling. Combining such product-level sourcing data with Nielsen Retail Scanner product price data (discussed shortly), we confirm that products accredited for responsible sourcing mainly fall in the specialty coffee category. Consistent with the literature, certified products are priced at \$9.8 on average, products with self-made claims are priced at \$9 on average, and the remaining products are priced at \$8 on average.<sup>6</sup> Figure 3 shows the categories' respective product counts and retail-price histograms.



**Figure 3 Product counts and prices by category of ESG labeling.**

Data summarized from Nielsen Label Insights and Retail Scanner data. Nielsen Label Insights was used to classify products based on ESG marketing on product packaging; the categories are products certified through a major 3rd party certifier (*e.g.*, Fair Trade), products making a self-made ESG claim, or none.

**2.2.2. Country-based sourcing.** We collect data on first-tier, cross-border suppliers from the Panjiva platform. Data are extracted for 41 of the 49 companies matched between Nielsen Label Insights, Retail Scanner, and RepRisk over the period from 2006 through 2016. Each company file contains information on every first-tier supplier that the company received an international shipment from, how many shipments were received, where the shipments came from, where the first-tier supplier is located, and the contents of the shipments. This information is used to develop a view of the countries each company is sourcing from, the percentage of shipments received from each country, and the diversity of country sourcing for each company based on first-tier suppliers. We manually filtered out all non-coffee shipments, *e.g.*, ancillary products such as straws or cups.

Figure 8 in Section 3 maps the source country of coffee shipments into the United States coming from coffee producing countries in our data. Shipments from North America and Europe are likely to originate from middlemen in the coffee supply chain, while shipments from South America,

<sup>6</sup> These trends are consistent when examining price per ounce of coffee; however, only half the products have information on product weight.

Africa and Asia are likely to be from the countries where the coffee was produced; the latter claim is supported by our observation that the aggregate mix of sourcing closely matches the mix of coffee production at the country level (International Coffee Organization 2019) and the continent level (World Bank 2015). Please see Appendix C for the comparison of country level coffee production against the sourcing data used in this study.

Our study's outcome variable is ESG violations that occur during coffee production. In our primary analysis of country-based sourcing, we account for brands' US-bound shipments received from the various coffee-producing countries in explaining their ESG violation rates. In other words, a brand's violation rate may be low (high) due to sourcing from a low-risk (high-risk) country. All our results are backed by robustness checks that further control for the brands' shipments received from middlemen countries (*e.g.*, located in North America and Europe) that do not produce coffee themselves.

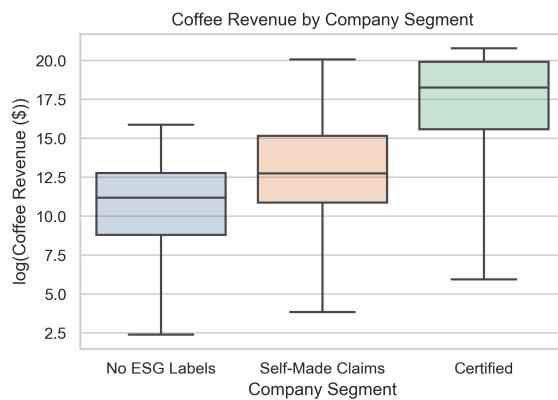
While Panjiva's visibility is limited to the first-tier of brands' cross-border supply chains, we believe that brands' shipments received directly from coffee-producing countries are highly likely to contain coffee grown within that country. The sourcing origin of the shipment's content is explicitly transcribed by approximately 25% of our 2,201 shipment records. For example, the record for a shipment originating in Colombia may specify in its contents field: "Colombian coffee beans". In 80% of records with such descriptions, coffee bean shipments originate from the country where the beans were grown. Only about 1% of such records identify shipment-origin and bean-origin countries differing in region or risk tier.

### 2.3. Control Variables

**2.3.1. Scale of production.** To appropriately interpret brands' ESG violation counts, we control for company-level production scale. Thus, we account for the fact that, *e.g.*, companies producing at a larger scale accumulate higher violation counts, even when keeping their violations *per sale* low. Therefore, we collect product-level sales data from Nielsen Retail Scanner and then aggregate to obtain company-level sales and revenue.

Figure 4 plots companies' coffee revenues in each of the three company categories. Companies larger by sales and revenues tend to have at least one product line with an ESG claim, whether third-party certified or self-made.

**2.3.2. News visibility.** RepRisk compiles brands' violation events when they are detected by news and regulatory sources, and the rates of detection likely vary with brands' news exposure levels. Thus, among a pair of companies incurring hypothetically the same violations, the company subject to higher news volume and scrutiny likely sees a greater number of its incidents brought to light and tallied. To measure a company's level of news exposure, we compute and use the volume of general news published about the company during the year of interest collected from LexisNexis.



**Figure 4 Boxplot of company-level coffee revenue by category of ESG labeling.**

Data from Nielsen Label Insights and Retail Scanner data from 2006-2016. Each data point is a company's coffee revenue in a given year. Companies are categorized based on having 3%+ of their product sales with third-party certification, remaining companies with some self-made claims on ESG sourcing, and companies with no ESG claims.

**2.3.3. Specialty/Commodity Share of Production.** Coffee companies have different product portfolio mixes of specialty versus commodity goods. The supply chains for companies that focus on specialty goods can look different than the supply chains of companies that primarily sell commodity coffee. To control for this choice in our robustness checks, we add controls of the average price in a company's product portfolio. Companies offering primarily specialty goods tend to have a higher average price across their portfolio.

## 2.4. Matching Datasets

The 6 datasets are merged using a fuzzy matching algorithm on company names to account for differences in naming convention across datasets. Matches were reviewed manually to verify and improve the quality of the connection. The matching process excluded many niche coffee products and small companies, but remained representative of the mainstream consumer coffee market. The final merged dataset pulls from a terabyte of data and includes 59% of coffee sales for 2016 with 1,827 coffee products, 130 brands, and 49 companies. Importantly, the dataset relevantly covers 89% of sales for products with claims and 93% of sales for products with endorsements, which are our main categories of interest. In the final sample, approximately 12% of the products in the merged dataset are third-party certified, while 15% have self-made claims.

## 3. Hypotheses

We describe stakeholders' concerns as well as the potential benefits of dual-sourcing and then develop our empirical hypotheses.

### 3.1. Stakeholders' Concerns about Dual-Sourcing

The prevalence of dual-sourcing strategies has raised concerns among NGOs, certifiers, and academics (Giovannucci and Koekoek 2003, Trudel and Cotte 2009, Thorlakson et al. 2018).

First, to the extent certification's goal is to achieve industry-wide changes in supplier behavior, dual-sourcing is potentially troubling. By catering to the subset of socially responsible consumers through certified goods while supplying questionably sourced goods to the rest of the market, dual-sourcing may detrimentally segment the market. Empirically, the market shares of certified products often suspiciously echo the shares delivered into the world's wealthiest, eco-conscious consumer markets. For example, the Roundtable on Sustainable Palm Oil (RSPO) reports certifying 15% of global palm oil in 2012–2013, while US and Europe—the source of nearly all of the demand for sustainable palm oil—consumes roughly 13–14% of the global supply (Waldman and Kerr 2014). Raising similar concerns, we find that dual-sourcing US coffee brands placed ESG-related claims on only 17% of their uncertified products. Empirical studies have yet to confirm the degree to which dual-sourcing brands' uncertified products harbor poor supplier behavior.

Second, the lack of transparency into dual-sourcing companies' supply chains may exert a negative externality by diluting the reputational credibility of certification labels such as Fair Trade. When a dual-sourcing firm is revealed to have ethical issues in its broader supply base, the reputational fallout may affect consumers' perceptions of the certifier's efficacy. A prominent coffee certification label's lead on chain of custody raised this concern in their conversation with us:

*"We don't know what companies are doing to source their non-certified products when they have a mix, it is a significant concern of ours... because it impacts the branding of certification."*

Thus, dual-sourcing may represent troubling risks for consumers and certifiers alike.

### 3.2. Potential Benefits from Dual-Sourcing

On the other hand, both anecdotal case studies and our own first-hand conversations with stakeholders suggest that a firm's certified production can generate positive spillover effects for its non-certified production.

First, certification programs facilitate finding and establishing reliable supplier relationships that are useful beyond certified production. The lead in charge of developing and implementing certification standards for a popular certification program noted to us that:

*"If a socially responsible retailer wants to source from a specific country, they just have to have the right relationships to know what to do. Often, we [the certifier] can help with that."*

In our first-hand interviews, the supply chain leads of several dual-sourcing coffee retailers noted that they source both certified and non-certified goods from the same suppliers or middlemen because they had learned which ones to trust through their certified production. Anecdotal case studies support positive spillovers from successfully cultivating good supplier relationships (Boström 2015).

Second, piloting a few products through certification may equip coffee brands with transferable operational knowledge regarding responsible sourcing. In a collection of anecdotal evaluations,

Barry et al. (2012) find that certification acts as a “laboratory for learning and demonstrating best practices,” ultimately leading to numerous practices adopted in non-certified sourcing channels. These include better management practices, cleaner technologies, and adherence to worker safety standards (Rueda et al. 2017, Locke et al. 2007b, Boudreau 2020). Once firms have developed the requisite operational and institutional knowledge, they may enjoy diminished marginal costs when further expanding the scope of their responsible sourcing to their other suppliers. Exemplifying how a well-placed intermediary can facilitate knowledge transfer, Li & Fung’s Vendor Support Services unit works closely with apparel suppliers to improve their efficiency and ESG compliance based on operational know-how gleaned from the 15,000 suppliers in its network (Lee and Tang 2018).

Thus, in contrast to the concerns about dual-sourcing limiting ethical sourcing, dual-sourcing brands may actually convert their experiences from certifying products into ethical sourcing for the rest of their supply chains.

### 3.3. Empirical Questions

To shed light on dual-sourcing’s effects in practice, we address two basic empirical questions.

First, we examine the widespread concern: do dual-sourcing brands source less responsibly for their uncertified products? To address this concern, we investigate whether brands certifying less of their product sales exhibit heightened ESG violation rates. A brand’s excess violations from uncertified production could be conceptualized much like a tax: *i.e.*, the brand incurs excess violations in the form of a tax applied on the amount of production it sources from non-certified channels. As a dual-sourcing brand certifies more of its product line in terms of sales, it decreases the tax base for excess violations—hence reducing its number of excess violations over a fully certified brand:

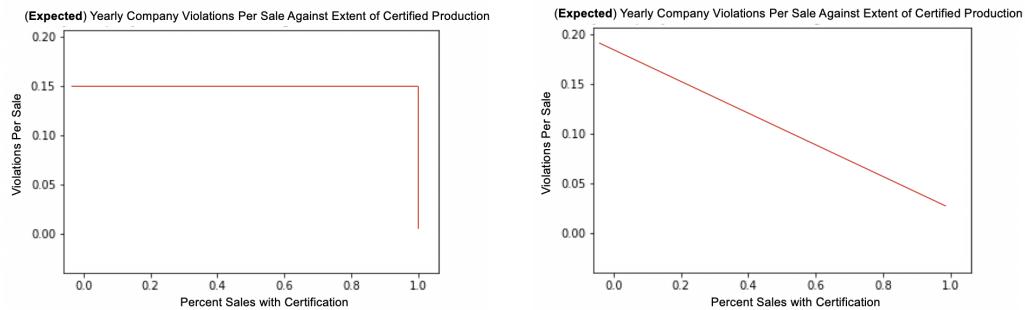
$$\text{Brand's expected excess violations} = \tau \times \text{Brand's uncertified product sales}, \quad (1)$$

where  $\tau$  represents the brand’s expected excess violations per unit when that production unit is uncertified instead of certified. The size of  $\tau$  informs us about the justifiable level of concern about coffee brands’ dual-sourcing. We empirically control for brand-level characteristics for a more accurate characterization.

Second, we test a key prediction of learning-based spillovers: do modest levels of product certification disproportionately improve brands’ overall ESG violation rates? When spillovers are potent, we expect dual-sourcing brands’ excess violations to dissipate quickly after they certify just a few products. This occurs for two reasons: (1) first, the tax base (*i.e.*, the brand’s uncertified sales) shrinks; and (2) second, the tax rate  $\tau$  (*i.e.*, the brand’s excess violations per sale over certified products) falls. The latter effect on  $\tau$ —where the brand learns to make its uncertified sourcing channels just as responsible as its certified sourcing channels, as  $\tau$  approaches zero—interests us empirically.

Two cases are commonplace in the literature, which does not consider spillover effects. At one extreme, the dual-sourcing brand faces reputational fallout whenever linked to a supplier found to have an ESG violation event, regardless of how much (or how little) it sources from the supplier (see, *e.g.*, discussion in Guo et al. 2015, Chen and Lee 2016). In this case, the brand's excess violations per sale are elevated and constant until it certifies its entire product line (see left panel of Figure 5).

In the second case, the brand's rate of excess violations per uncertified sale  $\tau$  is constant. As the brand's share of certified sourcing increases, the ESG violation risk decreases linearly (see right panel of Figure 5): *e.g.*, at 60% non-certified production, a brand doubles its excess violations rate over 30% non-certified production.

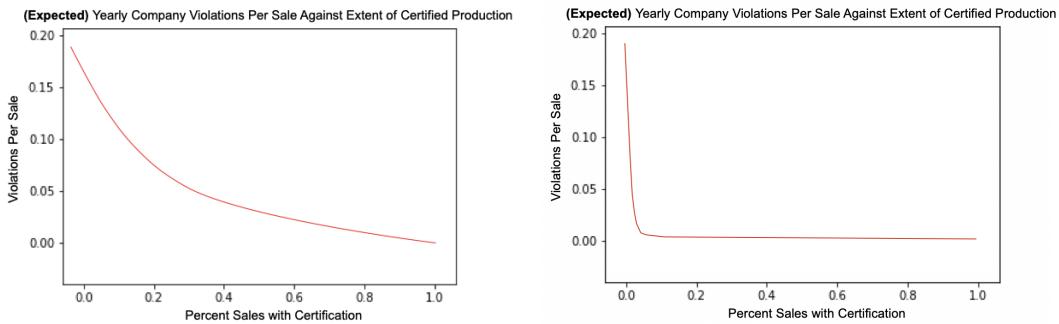


**Figure 5 Simulated excess ESG violations per sale against percentage of certified sales without spillovers. The left panel considers the case where any uncertified sourcing exposes the firm to equal ESG risk; the right panel considers ESG risk that scales linearly with the extent of uncertified sourcing.**

In the absence of spillover effects, both intuition and the literature suggest that the true relationship lies between the two canonical cases: excess ESG violation risk may attach rapidly upon first engaging non-certified suppliers and likely scales with the brand's degree of non-certified sourcing.

In contrast, when dual-sourcing generates positive, learning-based spillovers for the brand's uncertified supply channels, the excess violations rate  $\tau$  *decreases* as the brand's share of certified sourcing increases. Figure 5 simulates weak (left) and strong (right) spillover effects. The relationship can range from mildly-curved to a strong L-shaped drop as a function of the strength of the spillover effects.

We now turn to the data. To correspond to the example's plotted curves, Figure 7 plots detected ESG violations per sale by firm-year, against the firm's percentage of certified product sales. The plotted pattern is descriptive, pooling firms without controlling for their characteristics. The plot covers firm-years for 49 coffee companies between 2006 and 2016 and adjusts observations only by normalizing them against the time trend in yearly average violation rates.



**Figure 6 Simulated violations per sale against percentage of certified sales with spillovers. We consider weak (left) and strong (right) spillover effects.**



**Figure 7 Violations per sale against percentage of sales endorsed by company-year**

For 413 firm-years, we plot the violations per sale (y-axis) against the firm's share of sales from its certified products (x-axis). Violations per sale are mean-differenced by year to control for general time trends in violation rates.

The transparent plot reveals a surprisingly stark *L*-shaped relationship: firms with little to no certified sourcing post high violation rates, but the rates drop steeply for companies with some certified sourcing. Thus, contrary to the concerns of NGOs, certifiers and academics, dual-sourcing brands above a minor certification threshold largely fail to exhibit excess violations. Instead, a sharply *L*-shaped relationship reveals that violation rates are low once brands exceed the very modest threshold of 3% in percentage sales certified. With respect to ESG violations, companies that certify just 3% of their sales perform as well as companies that certify all of their sales. In Section 4.1's empirical analysis, we test for this relationship while additionally controlling for news-based violation detection rates and firm characteristics.

In Appendix E, we demonstrate a remarkably similar *L*-shaped relationship in the chocolate industry, suggesting that such learning-based spillovers are not unique to coffee; however, other industries currently have insufficient penetration of dual-sourcing strategies to carry out a careful econometric analysis compared to a mature market like coffee.

### 3.4. Sourcing Mechanism

In sourcing, brands face the critical decision of *where* to source from.

First, US coffee brands' international sourcing decisions vary. Figure 8's top panel shows the heat map of US-destined coffee shipments received from first-tier suppliers located in coffee-producing countries around the globe. On average, a US coffee brand sources directly from 8 coffee-producing countries and from 3 intermediate suppliers located in non-producing countries.

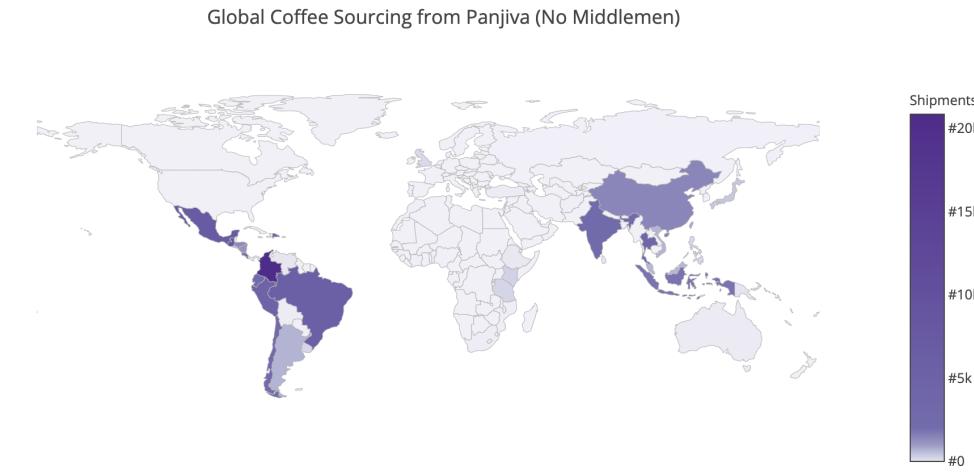
Second, the ESG risk attached to suppliers varies significantly as a geographic function of the source country: some countries are relatively safe (*e.g.*, Colombia), while others carry high risk of ESG violations (*e.g.*, Indonesia). Figure 8 contrasts the global heat map of US-destined coffee shipments from coffee-producing countries (top panel) against the countries' ESG violations per coffee shipment (bottom panel).

Underlying such variation, the adoption of certifications and sustainable-sourcing practices by farmers varies by country (Barry et al. 2012). On the difficulty of finding reliable suppliers in some parts of the world, the sourcing lead of a regional US-based coffee brand with ESG claims told us:

*"We have a mission to maintain ethical sourcing no matter what country we go to. We tried very hard to source from Indonesia but could not find suppliers that we trusted or mechanisms for making sure our suppliers were maintaining our standards, so we left."*

Thus, we distinguish between ethical sourcing that is accomplished from within "safe" countries versus by undertaking the difficult challenge of sourcing ethically from within high risk regions. The latter type of ethical sourcing contributes to the broader mission of social responsibility by strengthening emerging economies and creating infrastructure that improves conditions in countries with high supplier risk (Lee and Tang 2018).

In Section 4.2's empirical analysis, we account for firms' geographic sourcing decisions. On one hand, the analysis validates Section 4.1's examination of the L-shaped relationship by confirming that it holds after controlling for dual-sourcing brands' geographic sourcing. However, we find the distinction between sourcing strategy segments more interesting than the distinctions in certification levels within-segment. Companies making their own uncertified ESG claims achieve low violations by sourcing from low-risk countries, whereas dual-sourcing firms uniquely achieve low ESG violations even when sourcing from countries with high supplier risk. Consistent with the learning hypothesis, certifying a product may equip brands to source responsibly from countries with high supplier risk.



(a) Sourcing map of first-tier suppliers, excluding non coffee-producing countries.



(b) Country level ESG violations (extracted from RepRisk) per coffee shipment sent to companies in our study (computed from Panjiva).

**Figure 8** Panjiva data tracking origin country for shipments of coffee beans into US ports from 2006-2016.

#### 4. Empirical Analysis

This section details our empirical analyses based on our coffee dataset described in Section 2.

In Section 4.1, we empirically test for evidence that companies taking dual-sourcing strategies experience positive spillover effects from their certified sourcing onto their non-certified sourcing.

In Section 4.2, we estimate differences in violation rates for other company segments, including companies with self-made claims and companies with no ethical sourcing strategy. We further investigate the impact of country sourcing mix on violation outcomes for these company segments.

#### 4.1. Sourcing Spillovers

As noted in Section 2.1.1, our response variable is the yearly count of ESG violation incidents captured by RepRisk. As such, count models are the appropriate model family for regression (Cameron and Trivedi 1986). We primarily employ a zero-inflated negative binomial regression specification. This is because our data exhibits overdispersion (*i.e.*, the variance of the response variable is greater than its mean), suggesting that the negative binomial model is more appropriate than a Poisson model (which assumes the variance and mean of the response variable are equal). The negative binomial model relaxes the Poisson model by additionally fitting a shape parameter to model overdispersion (Cameron and Trivedi 2001). An additional challenge in our data is that there are an excessive number of zero responses for some firms — likely due to variations in detection likelihood of ESG violations — causing a deviation from the standard count model (Loeys et al. 2012). We account for this using zero-inflation, which fits an additional probability that the observed count in a particular year for a company is zero.

*Negative Binomial:* Denote our response variable, the yearly count of detected ESG violation for a firm  $i$  in year  $t$ , as  $C_{i,t}$ . Let  $\text{Year}_t$  be a yearly fixed effect to control for violation/detection trends over time; let  $\log(\text{Sales})_{i,t}$  be the logarithm of product sales volume of company  $i$  in year  $t$ , to control for the impact of company size on violation counts. We used the logarithm of production volume because violations are likely to scale with a proportional increase in production size rather than raw production size. However, we also tested alternative specifications with the raw production volume and had consistent results. Companies with high news volume and scrutiny are also likely to have a higher probability of violations being *detected*, leading to higher observed counts; thus, we control for  $(\text{News Count})_{i,t}$ , the volume of general news related to company  $i$  in year  $t$  obtained through LexisNexis. Now, let our key independent variable  $(\% \text{ Sales Cert})_{i,t}$  be the percentage of a company's product sales that are sourced through certified channels. As discussed earlier, Figure 7 suggests very different responses to certification levels above versus below 3%. To capture this, we interact our dependent variable with indicator variables  $\mathbb{1}_{>3\% \text{ Cert},i}$  and  $\mathbb{1}_{<3\% \text{ Cert},i}$ .<sup>7</sup> Our count process rate  $\lambda_{i,t}^{NB}$  is estimated using maximum likelihood estimation with the following specification:

$$\begin{aligned}\lambda_{i,t}^{NB} = & \exp(\alpha_0^{NB} + \alpha_1^{NB} (\% \text{ Sales Cert})_{i,t} \times \mathbb{1}_{>3\% \text{ Cert},i} + \alpha_2^{NB} (\% \text{ Sales Cert})_{i,t} \times \mathbb{1}_{<3\% \text{ Cert},i} \\ & + \alpha_3^{NB} \text{Year}_t + \alpha_4^{NB} \log(\text{Sales})_{i,t} + \alpha_5^{NB} (\text{News Count})_{i,t} + \varepsilon_{i,t}).\end{aligned}$$

<sup>7</sup> We tested alternative specifications — *e.g.*, fitting a non-interacted single intercept  $\mathbb{1}_{<3\% \text{ Cert},i}$  rather than  $(\% \text{ Sales Cert})_{i,t} \times \mathbb{1}_{<3\% \text{ Cert},i}$  — and found consistent results.

*Zero-Inflated Negative Binomial:* In a given year, a firm may have zero detected ESG violations for two reasons: (1) there may be no ESG violations that year, or (2) the company experienced violations but they were not detected. Let  $V_{i,t}$  be the true (unobserved) count of ESG violations, while  $C_{i,t}$  remains the detected count of violations. Additionally, let  $\mathcal{E}_{i,t}$  be the random event that there was a failure to detect ESG violations for firm  $i$  in year  $t$ , resulting in a  $C_{i,t} = 0$  observation. Under the complement event  $\mathcal{E}_{i,t}^c$ , we assume that the true (unobserved) violation count  $V_{i,t}$  follows a negative binomial distribution with count process rate  $\lambda_{i,t}^{ZI}$ . Then, we can write

$$\begin{aligned}\Pr(C_{i,t} = 0) &= \Pr(\mathcal{E}_{i,t}) + \Pr(V_{i,t} = 0 \mid \mathcal{E}_{i,t}^c) \cdot \Pr(\mathcal{E}_{i,t}^c), \\ \Pr(C_{i,t} = j) &= \Pr(V_{i,t} = j \mid \mathcal{E}_{i,t}^c) \cdot \Pr(\mathcal{E}_{i,t}^c).\end{aligned}$$

Before specifying the estimation of  $\lambda_{i,t}^{ZI}$ , we model the detection failure  $\mathcal{E}_{i,t}$ . As discussed earlier, larger companies with high news volume and scrutiny are likely to have a higher probability of being detected with a violation. Thus, we can express the probability of detection failure for firm  $i$  in year  $t$  as follows:

$$\Pr(\mathcal{E}_{i,t}) = \frac{\exp(W_{i,t}\gamma)}{1 + \exp(W_{i,t}\gamma)},$$

where we regress

$$W_{i,t}\gamma = \gamma_0(\text{News Count})_{i,t} + \gamma_1 \log(\text{Sales})_{i,t}.$$

We jointly estimate the count process rate  $\lambda_{i,t}^{ZI}$  for  $V_{i,t}$  using maximum likelihood estimation:

$$\begin{aligned}\lambda_{i,t}^{ZI} &= \exp(\alpha_0^{ZI} + \alpha_1^{ZI}(\% \text{ Sales Cert})_{i,t} \times \mathbb{1}_{>3\% \text{ Cert},i} + \alpha_2^{ZI}(\% \text{ Sales Cert})_{i,t} \times \mathbb{1}_{<3\% \text{ Cert},i} \\ &\quad + \alpha_3^{ZI}\text{Year}_t + \alpha_4^{ZI} \log(\text{Sales})_{i,t} + \varepsilon_{i,t}).\end{aligned}$$

Note that we no longer control for  $(\text{News Count})_{i,t}$  since it is already accounted for in the regression for the probability of detection failure  $\Pr(\mathcal{E}_{i,t})$ .

*Results:* The regressions results for the zero-inflated negative binomial model are given in Table 2. We additionally include results from the vanilla negative binomial model (which does not account for detection likelihood and the resulting excess zeros reflected in our data) as a robustness check. These results demonstrate that the *L*-shape is indeed statistically significant. In particular, the interaction term  $(\% \text{ Sales Cert})_{i,t} \times \mathbb{1}_{<3\% \text{ Cert},i}$  has a far larger coefficient than the interaction term  $(\% \text{ Sales Cert})_{i,t} \times \mathbb{1}_{>3\% \text{ Cert},i}$ , implying a steep drop in violations at the 3% certification cutoff followed by a slight downward slope.

A counterfactual analysis of the zero-inflated negative binomial (negative binomial) model suggests that a firm with 2% certified product sales can expect a 78% (64%) reduction in violations by sourcing 3% of its product sales through certified channels; in contrast, a firm with  $x\%$  certified

Outcome: Yearly ESG violations by company

Variable	(1) Zero-Inflated (ZINB)		(2) Negative Binomial (NB)	
	Estimate	Std Error	Estimate	Std Error
(Intercept)	1.80**	0.64	-2.99***	0.77
(% Sales Cert)* $\mathbb{1}_{>3\% \text{cert}}$	-3.91***	0.88	-2.94***	1.03
(% Sales Cert)* $\mathbb{1}_{<3\% \text{cert}}$	69.5**	22.6	47.3	30.2
Year	0.09	0.06	0.16**	0.06
Log(Sales)	-.02	0.04	0.22***	0.04
News Count			0.004***	0.0007
$\chi^2$ Goodness of Fit	270***		408***	

\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$

**Table 2 Zero-inflated negative binomial regression results for ESG violations by sourcing segment**

Zero-inflated negative binomial (ZINB) regression estimates using a panel of 413 firm-years. We use the chi-squared goodness of fit test for count-data models (Cameron and Trivedi 2010).

product sales, where  $x > 3$ , can only expect a 3.8% (2.1%) reduction in violations by sourcing  $(x + 1)\%$  of its product sales through certified channels. We test a number of additional regression specifications in Table 3, and find similar counterfactuals in all cases. In other words, 3% certification is sufficient<sup>8</sup> to achieve nearly all of the ESG-related benefits associated with certifications, and additional certification throughout the supply chain provides diminishing returns with respect to violation counts. These results are consistent when using a 5% certification threshold, suggesting that the result is robust to the choice of cutoff.

# Model	Controls	From 2% to 3% in Certified Sales	From $x\%$ to $(x + 1)\%$ for $x > 3$ in Certified Sales
1 NB	Year, News Counts, Log(Sales)	-64%	-2.1%
2 ZINB	Year, News Counts, Log(Sales)	-78%	-3.8%
3 ZINB	Year, News Counts, Sales	-66%	-3.8%
4 ZINB	Year, News Counts, Log(Sales), Price Avg	-61%	-4.0%

**Table 3 Estimated effect on ESG violations when a brand certifies an additional 1% of sales.**

The impact of a 1% increase in certified sales on the average company's ESG violations is calculated using coefficient estimates from four alternative regression models. We define an average company as one with the average values of the controls in our empirical distribution; we then use our estimated model to predict the impact of increasing certification coverage to compute counterfactuals.

*Propensity Score Matching:* The choice to dual-source with 3+ % certified sales is made based on firm-specific characteristics and thus does not satisfy random assignment. For instance, companies with more funding and scrutiny have a higher probability of certifying products to avoid negative

<sup>8</sup> In our data, companies with positive but less than 3% certified sales tend to be large firms with a wide product variety. We hypothesize that they use certification of 1–2 niche products as a strategy to diversify product offerings than to implement responsible sourcing.

publicity from ESG violations. We attempted to allay this concern in our regression by controlling for key firm-specific characteristics, such as company size and news volume. We now employ propensity score matching to test the robustness of our results (Rosenbaum and Rubin 1983).

Our treatment variable  $T$  is a binary indicator for whether a firm sources more than 3% of its product sales through certified channels. In our primary specification, we estimate a treatment propensity score

$$\pi_i = \beta_0 + \beta_1(\text{Sales})_i + \beta_2(\text{News Count})_i + \beta_3(\text{Avg Price})_i + \varepsilon_i,$$

controlling for company size through coffee sales, media scrutiny through general news counts, and specialty vs. commodity mix through average price (see Section 2.3). We take our outcome variable to be the yearly ESG violation count normalized by product sales. We use standard propensity score matching on treated and untreated units with propensities between [.1,.9], and employ Abadie-Imbens standard errors (Abadie and Imbens 2016).

We estimate a reduction of  $-0.0085$  violations per product sale by adopting dual-sourcing ( $p = 0.04$ ); supporting the result that dual-sourcing companies have significantly lower ESG violations than their peers. We test two additional specifications in Appendix B as robustness checks; all specifications yield qualitatively similar results.

#### 4.2. CSR Ecosystem

We now expand our attention to include companies from other segments, namely companies with self-made claims and companies with no ethical sourcing claims. As noted in Section 3, we also study the extent to which violation rates can be explained by country-level sourcing choices. We do this through a series of zero-inflated negative binomial regression specifications for our response variable  $C_{i,t}$ , the yearly count of detected ESG violations for a firm  $i$  in year  $t$ .

*Across Company Segments:* First, we simply examine violations as a function of company segment. We use yearly fixed effects  $\text{Year}_t$  to control for violation/detection trends over time and product sales volume log( $\text{Sales}_{i,t}$ ) to control for company size. We use the indicator variables  $\mathbb{1}_{>3\% \text{ Cert},i}$  and  $\mathbb{1}_{\text{Claims},i}$ , to denote if firm  $i$  has 3+% certified sales and is making self-made ethical sourcing claims respectively. (Note that the baseline firm makes no ESG-related claims.) Our count process rate  $\lambda_{i,t}^A$  is estimated using maximum likelihood estimation with the following specification:

$$\lambda_{i,t}^A = \exp(\alpha_0^A + \alpha_1^A \mathbb{1}_{>3\% \text{ Cert},i} + \alpha_2^A \mathbb{1}_{\text{Claims},i} + \alpha_3^A \text{Year}_t + \alpha_4^A \log(\text{Sales})_{i,t} + \varepsilon_{i,t}).$$

*Country Sourcing:* Next, we investigate the impact of different country sourcing strategies on violation outcomes. As discussed in the previous section, violations tend to be much higher when sourcing from certain countries (see Figure 8). Thus, we add controls for the share of coffee sourced

from each of the top 17 coffee-producing countries (these countries alone make up 95% of global coffee production (International Coffee Organization 2019); see list in Appendix D). Specifically,  $(\text{Coffee Producing Country})_{i,c}$  denotes the fraction of firm  $i$ 's coffee shipments that arrive from country  $c$  as noted in our Panjiva data. Our count process rate  $\lambda_{i,t}^B$  is estimated using maximum likelihood estimation with the following specification:

$$\begin{aligned}\lambda_{i,t}^B = & \exp(\alpha_0^B + \alpha_1^B \mathbb{1}_{>3\% \text{Cert}_i} + \alpha_2^B \mathbb{1}_{\text{Claims},i} + \alpha_3^B \text{Year}_t + \alpha_4^B \log(\text{Sales})_{i,t} \\ & + \sum_{c=1}^{17} \beta_c^B (\text{Coffee Producing Country})_{i,c} + \varepsilon_{i,t}).\end{aligned}$$

The previous regression focuses only on *direct sourcing* strategies. As discussed earlier, the Panjiva data only allows us to study sourcing from first-tier suppliers; restricting to coffee-producing countries in the above regression ensures that we can identify the source of the beans (we manually validated this claim; see Section 2 for details). Yet, we may be concerned that omitting the frequency of indirect sourcing (*i.e.*, through the use of middlemen in non-coffee producing countries) may lead to omitted variable bias. Therefore, we run a robustness check where we include a covariate  $(\% \text{Middlemen Sourcing})_i$  that captures the total percentage of shipments by firm  $i$  that were sourced from middlemen. Furthermore, we modify our 17 country-specific fixed effects  $(\text{Top Sourcing Country})_{i,c}$  to include the top sources of coffee shipments (thereby including both middleman and coffee-producing countries). Our count process rate  $\lambda_{i,t}^C$  is estimated using maximum likelihood estimation with the following specification:

$$\begin{aligned}\lambda_{i,t}^C = & \exp(\alpha_0^C + \alpha_1^C \mathbb{1}_{>3\% \text{Cert}_i} + \alpha_2^C \mathbb{1}_{\text{Claims},i} + \alpha_3^C \text{Year}_t + \alpha_4^C \log(\text{Sales})_{i,t} \\ & + \alpha_5^C (\% \text{Middlemen Sourcing})_i + \sum_{c=1}^{17} \beta_c^C (\text{Top Sourcing Country})_{i,c} + \varepsilon_{i,t}).\end{aligned}$$

*Results:* The key takeaways for the regression results for all three specifications are given in Table 4 below (the full regression tables are given in Appendix D).

These results yield two important insights. First, the specification without sourcing demonstrates that both dual-sourcing companies with more than 3% certified sales and companies with self-made claims achieve significantly reduced ESG violation rates over companies with no claims. This stands in contrast to stakeholders' concerns that firms with self-made claims may simply be engaging in cheap talk and may not actually reduce social or environmental harm (Giovannucci and Koekoek 2003, Darnall and Sides 2008). In particular, dual-sourcing companies reduce ESG violations by an estimated 569% relative to firms with no claims, while companies with self-made claims reduce ESG violations by an estimated 278% relative to firms with no claims. These rates are indeed not the same, and a Wald test confirms that dual-sourcing companies with more than 3% certified sales

Variable	Outcome: Yearly ESG violations by company		
	(1) No Sourcing Controls	(2) Control for Direct Sourcing	(3) Control for Direct and Middleman Sourcing
(Intercept)	-0.71 (0.67)	-1.247 (-0.808)	0.75 (1.09)
log(Sales)	0.29*** (0.057)	0.17*** (0.04)	-0.0061 (0.047)
Year	0.20** (0.061)	0.17*** (0.039)	0.14*** (0.033)
Self-made Claims	-2.78*** (0.66)	-0.21 (0.43)	-0.89 (0.77)
Certified 3%+	-5.69*** (0.73)	-3.97*** (0.68)	-2.79** (1.00)
Middlemen Sourcing			162.0*** (20.8)
Sourcing from top 17 coffee producing countries		Yes	
Sourcing from top 17 countries (including middlemen)			Yes
$\chi^2$ Goodness of Fit	305***	228***	195***

\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$

**Table 4 Zero-inflated negative binomial regression results for ESG violations by sourcing segment, after controlling for brands' country-level sourcing and time trend.**

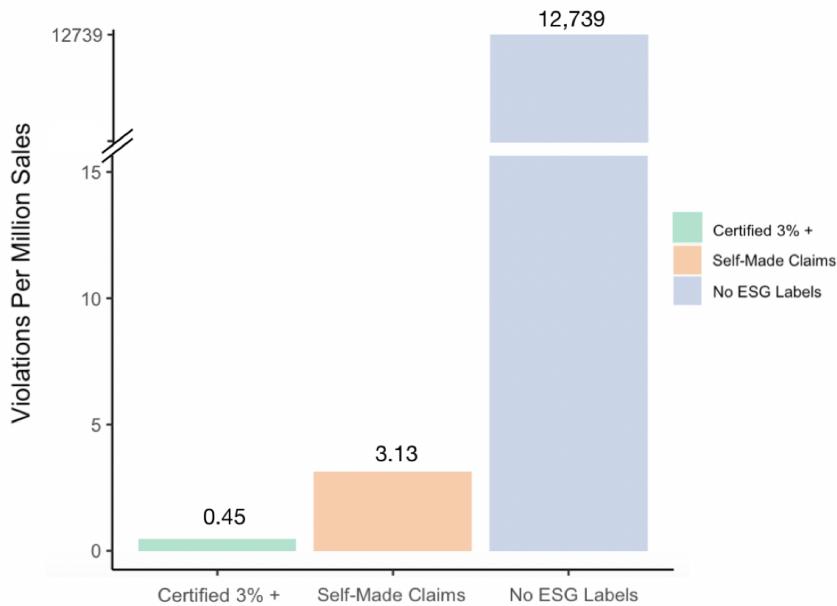
Zero-inflated negative binomial regressions using a panel of 355 firm-years. In column 2, we control for each brand's share of direct sourcing from each of the top 17 coffee-producing countries; in column 3, we control for each brand's share of sourcing from each of the top 17 countries by coffee shipments (which includes sourcing from intermediaries located in non-coffee producing countries). Column 3 additionally controls for brands' overall percentages of sourcing from middlemen countries. Our full results are provided in Appendix D. We use the chi-squared goodness of fit test for count-data models (Cameron and Trivedi 2010).

achieve significantly better violation reductions than companies with self-made claims ( $p < 0.001$ ). Still, these results indicate that stakeholders should be more concerned about the company segment with no claims, since their supply chains are the leading cause of global ESG violations.

Second, the specification controlling for sourcing from coffee producing countries shows that dual-sourcing companies with more than 3% certified sales continue to achieve significantly reduced ESG violations even after controlling for their sourcing mechanism; in contrast, the relatively low ESG violation rates of companies with self-made claims is largely explained by their sourcing mechanism. In particular, we see that sourcing strategies explain only 30% of the improved ESG performance of dual-sourcing companies, but account for 92% of the improved ESG performance of companies with self-made claims. Our third specification, which additionally controls for middleman sourcing, corroborates these results. Therefore, while companies with self-made claims seem to produce ESG outcomes that are roughly in line with the standard ESG performance in those countries, dual-sourcing companies seem to bring new methods to the table (e.g., possibly through better management practices, reliable relationships, experience with setting standards, etc.).

Underscoring this point, our data shows that when companies with more than 3% certified sales source from high risk countries, they are able to keep their violation rates 85% lower than companies with self-made claims (see Figure 9).<sup>9</sup> In other words, dual-sourcing companies appear to be creating infrastructure that allows responsible sourcing even in the challenging high risk regions identified in Figure 8.

Average Violations Per Million Sales for Companies Sourcing from High Risk Countries



**Figure 9 Average violations per million sales, shown for each sourcing segment's brands sourcing from high risk countries**

29 companies source directly from “high risk countries”, which are defined as the 10 countries with the highest ESG violations using all RepRisk ESG categories over the period 2006 to 2016. For each sourcing segment, we show the average of such companies’ violations per million sales. The segment averages shown are not adjusted for the constituent companies’ shares of sourcing from the high risk countries.

Unfortunately, despite these evident advantages, dual-sourcing companies source *less* from high risk countries than the other two company segments. As shown in Table 5, dual-sourcing companies source an average of 27% of their shipments from high risk countries, while companies with self-made claims source 38% and companies with no claims source 58% of their production from high risk regions. Based on our conversations with sourcing leads, this discrepancy may arise because

<sup>9</sup> We define “high risk” countries as the top ten countries with the highest ESG violations accumulated across all industries according to the RepRisk data used in this study. We find consistent results across alternative definitions of “high risk” countries, *e.g.*, countries with high ESG violations per capita, or limiting to coffee-specific ESG violations.

companies that cater to socially-conscious customers may wish to avoid high-risk countries due to the limited infrastructure available for sourcing ethically from these regions (they are often faced with numerous challenges including more lenient regulatory environments, corruption, and more frequent usage of middlemen that prevent transparency in the supply chain). Yet, our results show that these are the same companies that may be able to enable responsible sourcing in these regions.

	Mean	Std	Min	25%	50%	75%	Max
<b>Certified 3%+</b>	27%	40%	0%	0%	0%	36%	100%
<b>Self-Made Claims</b>	38%	36%	0%	0.3%	31%	73%	100%
<b>No ESG Labels</b>	58%	43%	0%	9%	85%	97%	99%

**Table 5 Brand-level shares of sourcing from high risk countries, as distributed within each sourcing segment**

We define “high risk countries” as the 10 countries with the highest total ESG violations using all RepRisk ESG categories over the period 2006 to 2016. A company’s share of sourcing from high risk countries is calculated as the fraction of shipments it received from high risk countries out of all sourcing shipments it received.

## 5. Discussion & Concluding Remarks

This paper highlights a novel role of certification in enabling companies to overcome the initial barriers of responsible sourcing, rather than a blanket solution intended to cover entire supply chains. Our primary finding is that certifying brands’ rates of CSR violations (adjusted for disparities in production volume and detection) are similarly low regardless of whether the brand’s portfolio is 5% certified or 100% certified. This motivates our view of certification-based responsible sourcing as a relatively *binary* outcome (we do not observe partially certified firms with “moderate” violations). In other words, taking a *first step* towards responsible sourcing by certifying a small portion of the product portfolio is a key first-order challenge. Our data suggests that it is not necessary for firms to invest in 100% certified product portfolios to achieve responsible sourcing (in contrast to, *e.g.*, Guo et al. 2015, Chen and Lee 2016).

This finding is important since our primary interviews with leaders of certified coffee supply chains suggest that certifying the entire product portfolio is an onerous task, and may actually hurt responsible sourcing outcomes by limiting entry into some emerging economies. Current certification standards are typically applied globally; however, it is well-recognized that differences in local culture and labor markets can make inflexible standards difficult to adopt in certain regions (Barry et al. 2012, Boström 2015). For example, the supply chain lead at a UK coffee company told us,

“We need exceptions based on specific country conditions, for example in India the agricultural labor standard in most certifications would require law changes that would either be extremely

*expensive and/or difficult to change. In addition, housing and living conditions in the standards do not account for cultural differences in different countries, for example housing capacity limits [intended to prevent unsafe or exploitative crowding in worker living conditions] are too small for Indian families that want to live together.”*

As discussed in Section 3.4, another sourcing lead described challenges in finding responsible suppliers in “risky” countries such as Indonesia. The lack of flexibility in current certification standards greatly limits the ability for socially-conscious firms to source from emerging economies in their certified production. Thus, given that a primary goal of responsible sourcing is to support welfare improvements in emerging economies—*e.g.*, sourcing by responsible firms has been shown to improve incomes, education rates, and regulatory governance (Dragusanu and Nunn 2018)—pushing for blanket certification of supply chains may *hinder* rather than support this goal.

Fortunately, our country-based analysis suggests that dual sourcing firms are able to leverage learning-based spillovers to responsibly source from “risky” countries despite the lack of infrastructure for doing so. This result motivates the view of certifications as “learning labs,” allowing firms to learn how to create and manage reliable supplier relationships while circumventing the inflexibility of current certification standards. This is an important contribution on the part of dual-sourcing firms both in the short term (achieving low ESG violations while improving welfare in risky countries), and the long term (in many cases, local governments have adopted voluntary standards set by these firms (Barry et al. 2012)). Thus, a promising agenda for future research is designing appropriate incentives that encourage dual-sourcing firms to engage emerging economies, and create the requisite infrastructure for ethical sourcing.

## Acknowledgments

The authors gratefully acknowledge research funding from the Wharton Social Impact Initiative and research assistance from Angel Chung and Avideep Pradhan. We thank Vishal Agrawal, Saed Alizamir, Noah Gans, Karen Zheng, and seminar participants at the 2019 INFORMS Annual Meeting, the Wharton School, and the 2020 Early-career Sustainable OM Workshop for providing valuable feedback.

Our analyses are derived based in part on data from The Nielsen Company (US), LLC and marketing databases provided through the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are our own and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

## References

- Abadie, Alberto, Guido W Imbens. 2016. Matching on the estimated propensity score. *Econometrica* **84**(2) 781–807.
- Abbasi, Maisam. 2017. Towards socially sustainable supply chains – themes and challenges. *Eur. Business Rev.* **29**(3) 261–303.

- Akkaya, Duygu, Kostas Bimpikis, Hau Lee. 2020. Government interventions to promote agricultural innovation. *Manufacturing & Service Oper. Management*.
- Amengual, Matthew, Greg Distelhorst. 2019. Can sourcing help enforce global labor standards? Evidence from the Gap Inc supply chain. *Evidence from the Gap Inc Supply Chain (October 9)*.
- Anagnostou, Angela, Paul TM Ingenbleek, Hans CM van Trijp. 2015. Sustainability labelling as a challenge to legitimacy: Spillover effects of organic Fairtrade coffee on consumer perceptions of mainstream products and retailers. *J. Consumer Marketing* **32**(6) 422–431.
- Babich, Volodymyr, Christopher S Tang. 2012. Managing opportunistic supplier product adulteration: Deferred payments, inspection, and combined mechanisms. *Manufacturing & Service Oper. Management* **14**(2) 301–314.
- Barry, Mike, Ben Cashore, Jason Clay, et al. 2012. Toward sustainability: The roles and limitations of certification. *RESOLVE*.
- Bastani, Hamsa, Joann F de Zegher. 2019. Do policies with limited enforcement reduce harm? Evidence from transshipment bans. *SSRN Working Paper*.
- Bloom, Nicholas, Christos Genakos, Ralf Martin, Raffaella Sadun. 2010. Modern management: good for the environment or just hot air? *The Economic Journal* **120**(544) 551–572.
- Boström, Magnus. 2015. Between monitoring and trust: commitment to extended upstream responsibility. *Journal of business ethics* **131**(1) 239–255.
- Boudreau, Laura. 2020. Multinational enforcement of labor law: Experimental evidence from bangladeshs apparel sector. *Working paper*.
- Buell, Ryan W, Shwetha Mariadassou, Yanchong Zheng. 2019. Relative performance transparency: Effects on sustainable choices. *Harvard Business School Technology & Operations Mgt. Unit Working Paper* (19-079).
- Cameron, A Colin, Pravin K Trivedi. 1986. Econometric models based on count data. Comparisons and applications of some estimators and tests. *J. Appl. Econom.* **1**(1) 29–53.
- Cameron, A Colin, Pravin K Trivedi. 2001. Essentials of count data regression. *A Companion to Theoretical Econometrics* **331**.
- Cameron, A Colin, Pravin K Trivedi. 2010. Microeconometrics using stata (revised ed.). *Number musr in Stata Press books. StataCorp LP*.
- Caro, Felipe, Leonard Lane, Anna Sez de Tejada Cuenca. 2018. Can brands claim ignorance? Unauthorized subcontracting in apparel supply chains. *UCLA Anderson working paper*.
- Castka, Pavel, Charles J Corbett, et al. 2015. Management systems standards: diffusion, impact and governance of iso 9000, iso 14000, and other management standards. *Foundations and Trends® in Technology, Information and Operations Management* **7**(3–4) 161–379.
- Chen, Jiayu, Anyan Qi, Milind Dawande. 2019. Supplier centrality and auditing priority in socially-responsible supply chains. *Manufacturing & Service Oper. Management* (forthcoming).
- Chen, Li, Hau L Lee. 2016. Sourcing under supplier responsibility risk: The effects of certification, audit, and contingency payment. *Management Sci.* **63**(9) 2795–2812.
- Chen, Shi, Qinjin Zhang, Yong-Pin Zhou. 2018. Impact of supply chain transparency on sustainability under NGO scrutiny. *Prod. Oper. Management*.
- Cho, Soo-Haeng, Xin Fang, Sridhar Tayur, Ying Xu. 2019. Combating child labor: Incentives and information disclosure in global supply chains. *Manufacturing & Service Oper. Management*.
- Consortium, The Sustainability. 2016. Greening global supply chains: From blindspots to actions). URL <https://www.sustainabilityconsortium.org/wp-content/themes/enfold-child/assets/pdf/2016-impact-report.pdf>.
- Corbett, Charles J, María J Montes-Sancho, David A Kirsch. 2005. The financial impact of ISO 9000 certification in the United States: An empirical analysis. *Management Sci.* **51**(7) 1046–1059.

- Darnall, Nicole, Stephen Sides. 2008. Assessing the performance of voluntary environmental programs: Does certification matter? *Policy Stud. J.* **36**(1) 95–117.
- de Zegher, Joann F, Dan A Iancu, Erica Plambeck. 2018. Sustaining rainforests and smallholders by eliminating payment delay in a commodity supply chain—it takes a village. *Working paper*.
- Distelhorst, Greg, Richard M Locke, Timea Pal, Hiram Samel. 2015. Production goes global, compliance stays local: Private regulation in the global electronics industry. *Regulation & Governance* **9**(3) 224–242.
- Dragusanu, Raluca, Nathan Nunn. 2018. The effects of Fair Trade certification: Evidence from coffee producers in Costa Rica. Tech. rep., National Bureau of Economic Research.
- Fang, Xin, Soo-Haeng Cho. 2019. Cooperative approaches to managing social responsibility in a market with externalities. *Manufacturing & Service Oper. Management* (Forthcoming).
- Giovannucci, D., P. Liu, A. Byers. 2008. Adding value: Certified coffee trade in North America. *FAO: Value-adding Standards in the North American Food Market*.
- Giovannucci, Danielle, Freek Jan Koekoek. 2003. The state of sustainable coffee: A study of twelve major markets.
- Guo, Ruixue, Hau L Lee, Robert Swinney. 2015. Responsible sourcing in supply chains. *Management Sci.* **62**(9) 2722–2744.
- Guo, Xiaomeng, Guang Xiao, Fuqiang Zhang. 2017. Effect of consumer awareness on corporate social responsibility under asymmetric information. *Available at SSRN 3039862*.
- Hasle, Peter, Per Langaa Jensen. 2012. Ergonomics and sustainability—challenges from global supply chains. *Work* **41**(Supplement 1) 3906–3913.
- Huang, Lu, Jing-Sheng Jeannette Song, Robert Swinney. 2017. Managing social responsibility in multitier supply chains. *Available at SSRN 2837332*.
- ILO. 2018. Human rights in supply chains (human rights watch). URL <https://www.dol.gov/agencies/ilab/our-work/child-forced-labor-trafficking/child-labor-cocoa>.
- International Coffee Organization. 2019. International coffee organization: Trade statistics tables.
- Kalkanci, Basak, Erica L Plambeck. 2018a. Managing supplier social & environmental impacts with voluntary vs. mandatory disclosure to investors. *Georgia Tech Scheller College of Business Research Paper* (37) 16–5.
- Kalkanci, Basak, Erica L Plambeck. 2018b. Reveal the supplier list? a trade-off in capacity vs. responsibility. *Georgia Tech Scheller College of Business Research Paper* (18-4) 18–8.
- Kim, Sang-Hyun. 2015. Time to come clean? Disclosure and inspection policies for green production. *Oper. Res.* **63**(1) 1–20.
- Kraft, Tim, León Valdés, Yanchong Zheng. 2018. Supply chain visibility and social responsibility: Investigating consumers behaviors and motives. *Manufacturing & Service Operations Management* **20**(4) 617–636.
- Kraft, Tim, León Valdés, Yanchong Zheng. 2019. Motivating supplier social responsibility under incomplete visibility. *Manufacturing & Service Oper. Management* (forthcoming).
- Lee, Hau L, Christopher S Tang. 2018. Socially and environmentally responsible value chain innovations: New operations management research opportunities. *Management Sci.* **64**(3) 983–996.
- Levi, Retsef, Manoj Rajan, Somya Singhvi, Yanchong Zheng. 2019. Unifying agricultural wholesale markets: Impact on farmers' income. *Available at SSRN*.
- Lewis, Tracy R, Fang Liu, Jing-Sheng Song. 2019. A dynamic mechanism for achieving sustainable quality supply. *Fuqua School of Business, Duke University, Durham*.
- Li, Jun, Di Andrew Wu. 2017. Does corporate social responsibility benefit society? *Management Sci.* (forthcoming).
- Locke, Richard, Thomas Kochan, Monica Romis, Fei Qin. 2007a. Beyond corporate codes of conduct: Work organization and labour standards at Nike's suppliers. *Internat. Labour Rev.* **146**(1-2) 21–40.

- Locke, Richard M, Fei Qin, Alberto Brause. 2007b. Does monitoring improve labor standards? Lessons from Nike. *ILR Rev.* **61**(1) 3–31.
- Loeys, Tom, Beatrijs Moerkerke, Olivia De Smet, Ann Buysse. 2012. The analysis of zero-inflated count data: Beyond zero-inflated poisson regression. *British J. Math. Statist. Psych.* **65**(1) 163–180.
- Nelson, Valerie, Ximena Rueda, Walter JV Vermeulen. 2018. Challenges and opportunities for the sustainability transition in global trade (introduction). *Business Strategy and the Environment* **27** 173–178.
- Orsdemir, Adem, Bin Hu, Vinayak Deshpande. 2019. Ensuring corporate social and environmental responsibility through vertical integration and horizontal sourcing. *Manufacturing & Service Oper. Management* **21**(2) 417–434.
- Pagell, Mark, Zhaohui Wu. 2009. Building a more complete theory of sustainable supply chain management using case studies of 10 exemplars. *J. Supply Chain Management* **45**(2) 37–56.
- Plambeck, Erica L, Terry A Taylor. 2015. Supplier evasion of a buyer's audit: Implications for motivating supplier social and environmental responsibility. *Manufacturing & Service Oper. Management* **18**(2) 184–197.
- Porteous, Angharad H, Sonali V Rammohan, Hau L Lee. 2015. Carrots or sticks? Improving social and environmental compliance at suppliers through incentives and penalties. *Prod. Oper. Management* **24**(9) 1402–1413.
- Rosenbaum, Paul R, Donald B Rubin. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* **70**(1) 41–55.
- Rueda, Ximena, Rachael D Garrett, Eric F Lambin. 2017. Corporate investments in supply chain sustainability: Selecting instruments in the agri-food industry. *J. Cleaner Prod.* **142** 2480–2492.
- Sodhi, ManMohan S, Christopher S Tang. 2018. Research opportunities in supply chain transparency .
- The Guardian. 2011. How green is your coffee? URL <https://www.theguardian.com/environment/2011/oct/04/green-coffee>.
- Thorlakson, Tannis, Joann F de Zegher, Eric F Lambin. 2018. Companies' contribution to sustainability through global supply chains. *Proc. National Acad. Sci.* **115**(9) 2072–2077.
- Toffel, Michael W, Jodi L Short, Melissa Ouellet. 2015. Codes in context: How states, markets, and civil society shape adherence to global labor standards. *Regulation & Governance* **9**(3) 205–223.
- Trudel, Remi, June Cotte. 2009. Does it pay to be good? *MIT Sloan Management Rev.* **50**(2) 61.
- US Department of Labor. 2013. US Labor Department's initiative results in more than \$100,000 in back wages and penalties for violations affecting 150 workers in Hawaii's coffee industry. URL <https://www.dol.gov/whd/media/press/whdpressVB3.asp?pressdoc=Western/20130221.xml>.
- Waldman, Kurt B, John M Kerr. 2014. Limitations of certification and supply chain standards for environmental protection in commodity crop production .
- World Bank. 2015. Risk and finance in the coffee sector. *Agriculture Global Practice* .
- WWF. 2007. Gone in an instant URL [https://wwf.panda.org/wwf\\_news/?92080/](https://wwf.panda.org/wwf_news/?92080/).
- Zamora, Miguel. 2013. Farmworkers left behind: The human cost of coffee production. *Daily Coffee News* URL <https://dailycoffeenews.com/2013/07/17/farmworkers-left-behind-the-human-cost-of-coffee-production/>.
- Zhang, Han, Goker Aydin, Rodney P Parker. 2019. Social responsibility auditing in supply chain networks .

## Appendix

### A. Certification Labels

Table 6 includes the 12 unique certification endorsements on coffee products captured by the Nielsen LabelInsights data in 2016. On-pack certification labels were used to identify the percentage of products in a brand's portfolio that were produced under certified standards. Out of 4,534 products, 13% had formal certification endorsements on their packaging.

Coffee Certifications
Certified B Corporation
Certified SC Product
Columbian Coffee Growers Federation
Cooperative Coffees Member
Fair Trade Certified
Fair trade certified by Pareve
Fair Trade USA
Rainforest Alliance Certified
100% Rainforest Alliance Certified
30% Rainforest Alliance Certified
65% Rainforest Alliance Certified
Social Certified by IMO

**Table 6 Sustainability-related Third-party Certifications for Consumer Coffee Products**

## B. Propensity Score Matching

The analysis finds that dual-sourcing companies that source more than 3% of their portfolio from certified producers have similarly low ESG violations. This means that companies with only 3% certified sales perform as well as companies with 99% certified sales in regards to ESG violations. However, the choice to dual-source with more than 3% certification coverage is based on corporate strategy and is not randomly assigned. For example, if a company is large and has higher scrutiny, they may have a higher probability of certifying products to avoid negative publicity. We use propensity score matching to compare companies with similar attributes and test the robustness of our finding that companies with 3%+ certification have significantly lower ESG violation rates than their peers. We estimate the impact of the treatment variable of dual sourcing with more than 3% of the portfolio certified to analyze the effect on an outcome of ESG violations per sale. We match on metrics related to company size (*e.g.*, sales, revenue), scrutiny (*e.g.*, news counts), and product specialty (*e.g.*, average price). The results for the PSM are listed in the table below. All results show a significant treatment estimate, which supports our findings that companies with more than 3% certified sales have lower ESG violations than their peers.

Match Variables	Treatment Estimate	Std. Error
Sales, Revenue	-0.0037**	0.001
Sales, Revenue, News Counts	-0.019†	0.01
Sales, News Counts, Average Price	-0.0085*	0.004

† $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

**Table 7 Propensity-score matching estimates of the average treatment effect of sourcing 3+% certified sales for different sets of company controls.**

### C. Top Coffee Producing Nations (accounts for 95% of total production)

The Panjiva data used in our study leverages port documents in the United States to extract information on shipment recipient, supplier, and shipment contents. Therefore, Panjiva allows us to see the top tier suppliers that are shipping coffee beans directly to the United States. This table shows the level of coffee production sourced from the International Coffee Organization with the weight and share of shipments coming into the United States extracted from Panjiva. We see that the shipments received by the United States has an over-representation from South America relative to global coffee production. The United States may be receiving more shipments from South America due to convenience of proximity.

	Coffee Production (kg)	Production Share	US Shipment Weight (kg)	Shipment Share
Brazil	2,591,999,620	30%	600,805,606	18.2%
Vietnam	1,649,999,852	19%	25,554,904	0.8%
Colombia	809,999,853	9%	985,872,303	29.9%
Indonesia	659,999,577	8%	49,371,020	1.5%
Ethiopia	383,999,960	4%	3,539,833	0.1%
Honduras	347,999,695	4%	43,272,229	1.3%
India	347,999,695	4%	47,334,213	1.4%
Uganda	287,999,857	3%	216,214	0.0%
Mexico	233,999,685	3%	147,494,258	4.5%
Guatemala	203,999,993	2%	212,665,272	6.4%
Peru	191,999,753	2%	186,226,717	5.6%
Nicaragua	131,999,915	2%	51,819,000	1.6%
China	116,819,993	1%	32,843,933	1.0%
Cote d'Ivoire	107,999,889	1%	309,560	0.0%
Costa Rica	89,519,629	1%	412,355,396	12.5%
Kenya	49,979,982	1%	16,458,174	0.5%
Papua New Guinea	47,999,598	1%	2,021,298	0.1%

**Table 8 List of the top 17 coffee producing countries with level of production sourced from the International Coffee Organization and shipments to the United States sourced from the Panjiva data we used in our study.**

## D. Country Sourcing Regression Results

Below are the full regression results for the analysis of the impact of country sourcing strategies on violation outcomes. The first table includes the regression coefficients with controlling for sourcing from the top 17 coffee producing countries. The second table shows the regression coefficients with controlling for sourcing from the top 17 countries (including middlemen countries such as Italy) and a covariate for overall percentage sourced from middlemen countries by company. The results from both tables suggest that dual-sourcing companies with more than 3% certified sales achieve reduced ESG violations regardless of their sourcing strategy, whereas the low ESG violation rates for companies with self-made claims is primarily explained by their sourcing mechanism.

Outcome: Yearly ESG Violations by Company

Variable	Estimate	Std Error
log(Sales)	0.166***	(0.0442)
Year	0.167***	(0.0398)
Self-made Claims	-0.214	(0.425)
Certified 3%+	-3.967***	(0.676)
Ivory Coast	-427.7	(1388.8)
China	-29.09***	(7.066)
Colombia	-1.713	(0.946)
Ethiopia	-2295.4	(3583913.6)
Kenya	1109.2	(830430.2)
Uganda	0	(.)
India	4.177***	(0.834)
Indonesia	363.0***	(50.04)
Papua New Guinea	0	(.)
Vietnam	-173.7***	(25.06)
Costa Rica	21.33**	(8.234)
Guatemala	24.51	(15.53)
Honduras	429.4	(715.6)
Mexico	0.485	(0.512)
Nicaragua	-1349.6	(1841.4)
Brazil	-2.545***	(0.578)
Peru	-9.183	(6.570)
(Intercept)	-1.247	(0.808)

\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$

**Table 9 Zero-inflated negative binomial regression results for yearly ESG violations across company segments with controls for direct sourcing mechanism (origin of first-tier coffee shipments) for 355 firm-years.**

## Outcome: Yearly ESG Violations by Company

Variable	Estimate	Std Error
log(Sales)	-0.00613	(0.0467)
Year	0.137***	(0.0328)
Self-made Claims	-0.885	(0.768)
Certified 3%+	-2.785**	(1.006)
Middlemen Sourcing	162.0***	(20.78)
Colombia	3.639***	(0.974)
Guatemala	-12.94**	(4.219)
Italy	-187.2***	(22.35)
Mexico	-25.58***	(7.242)
United States	-166.2***	(21.55)
Spain	49.59	(36.33)
Brazil	-1.825**	(0.680)
Peru	-43.82***	(6.027)
Dominican Republic	-239.2*	(112.0)
Thailand	1.999	(107.5)
Germany	-236.6***	(29.79)
India	-0.437	(1.314)
Costa Rica	1.155	(12.29)
Chile	22.76	(22.50)
Ecuador	-2.818	(1.531)
Indonesia	-8.038	(9.518)
Switzerland	-160.9***	(38.74)
(Intercept)	0.750	(1.087)

\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$

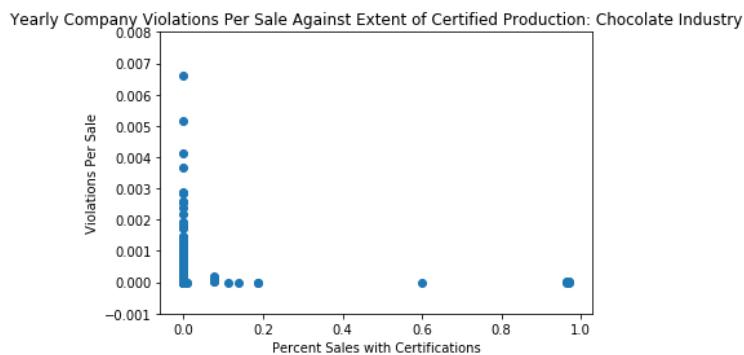
**Table 10 Zero-inflated negative binomial regression results for yearly ESG violations across company segments with controls for direct and middlemen sourcing mechanism (origin of first-tier coffee shipments) for 355 firm-years.**

## E. Data from Chocolate Industry

In order to understand whether similar spillover effects from certification occur in other industries, we performed the same data collection and merging process outlined in Section 2 for the chocolate industry (see Table 11 for a list of certifications for consumer chocolate products). In Figure 10, we plot the violations per sale for 100 chocolate companies against their percentage certified sales from 2006-2016. We see an *L*-shape pattern that is remarkably consistent with Figure 7 for the coffee industry. In particular, firms with little to no certified sourcing post high violation rates, but the rates drop steeply for companies with some certified sourcing. However, this sample only includes 9 dual-sourcing companies and 9 companies with self-made claims, while 82 companies make no ESG claims. Thus, we lack the data to econometrically verify this relationship while controlling for news-based violation detection rates and firm characteristics (as we did for the coffee industry, where many more firms adopted dual-sourcing strategies). However, the descriptive evidence is suggestive that dual-sourcing brands in the chocolate industry may also enjoy positive, knowledge-based spillovers encouraging responsible sourcing throughout their product lines.

<b>Chocolate Certifications</b>
100% Fair Trade Certified
Certified B Corporation
Certified Sustainable Palm Oil
Cocoa Life Certified
Direct Trade Certified
Fair for Life Certified
Fair Trade Certified
Fair Trade USA Certified
Fairwild Certified
Green Energy Certified
Heirloom Cacao Preservation Certified
Project Seal
Rainforest Alliance Certified
True Source Certified
Utz Certified
Utz Certified Good Inside
Worker Owned Coop

**Table 11 Sustainability-related Third-party Certifications for Consumer Chocolate Products**



**Figure 10 Chocolate industry violations per sale by percent sales with certification in 2016**

The chocolate industry shows a similar L-shape pattern when analyzing violation rates against percentage of sales with certification. The above graph includes 944 company-years for 100 companies from 2006-2016 (not all companies have products sold in every year).