Redacted Identities in Shipment Records: Evidence from Forced Labor Scrutiny in Supply Chains

Sandra G. Schafhäutle and Gurpal S. Sran*

February 2025
(Original draft: August 2024)

Abstract

US Customs and Border Protection allows firms to request redaction of their own (and their suppliers') identifying information in transaction-level shipment records. We find that about 16% of shipment records from 2013 through 2023 have redacted identities, with significant variation across time, origin regions, and other shipment characteristics. Along with examining proprietary costs as a motive for firms to redact identities, we focus on an important but understudied force: supply chain scrutiny costs related to forced labor risks. Consistent with such costs, shipments from countries with forced labor vulnerabilities and weak government responses to forced labor are more likely to have redacted identities. We then exploit a series of events related to forced labor allegations in international cotton and apparel production that intensified supply chain scrutiny related to forced labor risks. Using a difference-in-difference-in-differences design, we find an increase in redactions for affected cotton and apparel shipments after these events. Overall, our evidence suggests that importers redact identities from shipment records in the presence of supply chain scrutiny costs introduced by public and regulatory attention to corporate social responsibility.

Keywords: disclosure, redactions, supply chains, forced labor
JEL Classifications: M41, M48, F14, M14, J80, G38

*Schafhäutle: The Wharton School of the University of Pennsylvania (schaf@wharton.upenn.edu) and Sran: New York University Stern School of Business (gsran@stern.nyu.edu). We thank seminar participants at the 2024 Dartmouth Tuck Fall Accounting Workshop, Emerging Scholars Brownbag Series, Wharton, Rutgers, NYU, Baruch, and Fordham for helpful comments and suggestions. We also thank Yufei Chen for excellent research assistance. We gratefully acknowledge financial support from the Global Initiatives and ESG Initiative at Wharton, the Wharton School, and NYU Stern.

1. Introduction

Firms face increasing pressure from governments, investors, NGOs, and other stakeholders to enhance supply chain transparency.[1] Despite these growing calls for transparency and the possible signaling benefits of providing such disclosure, many firms refrain from comprehensively disclosing detailed information about their international supplier relationships, implying that they perceive material costs related to such disclosures. As a result of this supply chain opacity, external stakeholders face challenges in effectively monitoring supply chains, and academics are limited in their ability to study the forces that shape global supply chains and the transparency thereof.

In this paper, we study the characteristics of firms' observable *non-disclosure* decisions about their global supply chains. To do so, we leverage features of US import shipment records. According to US law and the Freedom of Information Act, Customs and Border Protection (CBP) is required to make data from maritime shipments publicly available, including the description of goods, quantity, and, if not redacted, identifying information of importers and suppliers (19 C.F.R. § 103.31).[2] To then measure firms' non-disclosure decisions, we exploit the fact that CBP allows firms to seek *manifest confidentiality*. This program allows importers to conceal their own names (and variants thereof), along with the identifying information of all their suppliers, from shipment records.[3] Specifically, upon receiving a manifest confidentiality application from an importer, CBP begins redacting identifying information from the importer's shipment records on a rolling basis for a two-year period such that these identities are not visible in public records. These records remain redacted indefinitely. After the two-year period, CBP allows for redaction renewal requests.

[1] Broadly, supply chain transparency refers to firms knowing where and how their goods are sourced and, in turn, *communicating* this information with their stakeholders (Bateman and Bonanni [2019]).

[2] CBP is the largest federal enforcement arm of the Department of Homeland Security (DHS).

[3] Similar data confidentiality provisions exist for outward (export) manifests in the US. We focus on inward (import) manifests given their relation to our research question. Because data on air-, truck-, and rail-based imports are not publicly available, we use data on maritime shipments. Note that maritime shipments account for about 65% of US imports measured by weight and about 45% of US imports measured by value, highlighting the importance of maritime trade data (see www.trade.gov).

Although manifest confidentiality complicates the study of non-disclosure decisions at the firm level (as redactions, by definition, conceal firms' identifying information), the remaining *transaction-level* data remains publicly observable and allows us to draw inferences about the forces that relate to such non-disclosure decisions.[4] Specifically, public shipment records still include information on the shipment's arrival date, the type and weight of goods shipped, and the countries, regions, and ports related to that shipment, even if identities are redacted.

Notably, importers are *not* required to provide evidence that the disclosure of identities would have caused substantial harm to their competitive position to be granted manifest confidentiality (19 C.F.R. § 103.31.d.ii). Consequently, while redaction decisions could be driven by competitor-based proprietary costs (Wagenhofer [1990]) and such costs are anecdotally cited to motivate the manifest confidentiality program (Kienzle [2022]), we primarily focus on a significant but understudied force that could shape redaction decisions: supply chain scrutiny costs related to forced labor risks.[5]

Because the revelation of trade linkages exposed to forced labor risks (such as poor working conditions and child labor) can induce scrutiny, reputational damage, and boycotts (e.g., Baron [2001, 2003]; Egorovy and Harstadz [2017]), importers could be motivated to conceal such linkages from the public. Although CBP has access to all transaction data regardless of manifest confidentiality, the ability of external parties to observe and process such data remains a key input in CBP's ability to combat forced labor in supply chains (Koscak [2022]). As a result, importers' identity redactions through manifest confidentiality can be an effective strategy to minimize both public *and* regulatory scrutiny costs. Furthermore, given the growing concern about forced labor in production facilities worldwide (ILO [2024]) and US regulatory actions that have targeted forced labor risks in an accelerating fashion since the nineteenth century (CSR [2023]), an investigation of the forces that motivate importers'

[4]We lack information on which importers have requested confidentiality and the timing of their request. Consequently, our results in the paper reflect both (i) the redaction choices made by importers and (ii) the market share of redacting importers within specific supply chains over time.

[5]Throughout the paper, we use the terminology *forced labor* when referring to an individual being forced to provide work or services against their will or without voluntary consent through the use of physical and psychological factors, threats, fraud, or coercion (see, e.g., www.walkfree.org).

non-disclosure decisions related to forced labor risks is important.

We first provide descriptive evidence on the nature of shipments with redacted identities and document several key insights.[6] By analyzing over 104 million transaction-level shipment records from Panjiva (based on CBP maritime shipment data), we find that about 16% of shipments have redacted importer identities and the share of such importer redactions is increasing over time. Of these 16%, about 93% also have redacted supplier identities.[7] Redactions are observed for shipments originating from various global regions, with notably higher redaction rates from Eastern Asia, Micronesia, Northern Africa, South-Eastern Asia, and Western Asia, where each region has an average share of redacted shipments of more than 16%. Across US destination regions, the New York and Pacific regions receive shipments with the highest average redaction rates (about 17%). Finally, redacted shipments tend to be smaller, on average, both in terms of weight and dollar value.

Then, we connect to an extensive literature documenting that competitor-based proprietary costs present a driving force underlying firms' (non-)disclosure decisions (e.g., Ellis et al. [2012]; Glaeser [2018]; Li et al. [2018]). Based on prior research documenting that the manufacturing sector relies on opacity to safeguard proprietary information (Shackelford and Kindlon [2021]), we examine whether shipments of goods more exclusively related to the manufacturing sector (e.g., chemicals and plastics) are more likely to have redacted identities. Compared to shipments of non-durable goods (e.g., footwear and toys), we find that manufacturing shipments are more likely to have redacted identities. In addition, we argue that importers are relatively more concerned about proprietary costs when supply chain relationships are newly formed and therefore expect higher rates of redaction in new linkages. Exploiting the granularity of the transaction-level data and identifying linkages as a unique triplet consisting of (i) country of origin, (ii) destination region, and (iii) good type, we

[6] Throughout the paper, we use the term *shipment* to refer to each importer-supplier transaction involving the receipt of a specific set of goods, as documented on a bill of lading.

[7] Since the redaction request is made by the importer, our main outcome variable is an indicator variable set equal to one when at least the importer's identity is redacted. Results throughout the paper are similar when we instead set the indicator variable equal to one when both importer and supplier identities are redacted.

find that new linkages are more likely to have redacted identities. In sum, these results are consistent with the notion that proprietary cost concerns induce supply chain opacity.

In the core part of our paper, we study whether supply chain scrutiny costs related to forced labor risks are associated with redaction decisions. Utilizing country-level measures on forced labor developed by Walk Free—an international human rights group—we assess whether shipments originating from countries with high compared to low scores on forced labor risks are more likely to have redacted identities. We find that, compared to shipments originating from countries with low scores on forced labor vulnerability (and while controlling for good type, destination region, timing, value, and weight), shipments originating from countries with high scores on forced labor vulnerability are more likely to have redacted identities. Alternatively, using a measure that reflects the strength of governments' responses in addressing forced labor issues, we continue to find higher redaction rates in shipments originating from countries that have taken limited steps to address forced labor relative to countries that have taken stronger steps. Overall, our findings support the idea that revelations of forced labor risks can impose material costs on importers. To avoid scrutiny, importers sourcing goods from suppliers with possibly questionable labor practices are therefore more likely to redact their identities in shipment records.

To further support these findings, we exploit a series of key events that increased scrutiny of importers' human rights due diligence in supply chains. Specifically, in March 2020, the Australian Strategic Policy Institute (ASPI), a non-partisan think tank, released a report that systematically linked forced labor claims—stemming from the alleged mass transfer of Uyghurs and other ethnic minorities to factories across China—to the supply chains of major importers (ASPI [2020]). Subsequently, CBP introduced measures that targeted all importers sourcing certain goods from China (primarily cotton and related products such as apparel), culminating with the implementation of the US Uyghur Forced Labor Prevention Act (UFLPA) in June 2022. The UFLPA establishes a rebuttable presumption of forced labor in certain supply chains and thereby increases the entry barriers for goods suspected to be produced by forced labor. We argue that this series of events represents a swift increase in supply chain scrutiny for US importers sourcing certain goods from China, therefore inducing changes in incentives related to supply chain (non-)disclosure behavior. Although we recognize that international sentiment is mixed regarding the extent to which forced labor conditions exist in Chinese production facilities (e.g., Xinhua [2020]; Farge [2022]), our conjecture is based on the notion that US importers sourcing goods from Chinese cotton and apparel suppliers perceive increased supply chain scrutiny linked to forced labor risks.

We exploit the timing of these events and study changes in redacted identities in cotton and apparel shipments from China using a difference-in-difference-in-differences design. Specifically, we define four cells: (i) cotton and apparel shipments from China, (ii) cotton and apparel shipments from elsewhere in the world, (iii) manufacturing- and durable-good shipments from China, and (iv) manufacturing- and durable-good shipments from elsewhere in the world. Because a difference-in-differences design between (i) and (ii) could simply reflect the endogenous effects of concurrent events on the redaction rate for any shipments originating from China (rather than cotton and apparel shipments specifically), we exploit the difference-in-differences over time between (iii) and (iv) to control for overall shocks to redaction rates in China-origin shipments.[8]

Estimating this difference-in-difference-in-differences design with three-way fixed effects over the period from 2018 to 2023, we find that following ASPI's report, there is a notable increase in the probability of redacted identities in cotton and apparel shipments from China. Estimating the effects in event time, we observe an initial spike in 2020, followed by a gradual attenuation of the effect in the latter half of 2021. After the implementation of the UFLPA in June 2022, the probability of redacted cotton and apparel shipments from China rose again through the end of 2023. Supplementary analysis indicates that this increase in redacted identities is at least in part due to new redaction requests. Overall, these findings

[8] Concurrent events include the "trade war" between the US and China in 2018 and 2019 (Fajgelbaum et al. [2024]), the COVID-19 pandemic that led to supply chain disruptions in 2020 (Liu et al. [2023]), and CBP making the manifest confidentiality request system available as an online form in 2020.

are consistent with the idea that importers redact identities to lower the expected costs associated with forced labor scrutiny by third parties.

Our paper makes three contributions. First, we add to the existing literature on the various motivations that drive firms' redaction decisions. Extensive prior research focuses on the role of proprietary costs as an important factor that influences redaction decisions in disclosures such as material contracts (e.g., Verrecchia and Weber [2006]; Glaeser [2018]; Chen et al. [2022]) or major customer lists (e.g., Ellis et al. [2012]; Li et al. [2018]). We add to this literature by providing evidence on another force that can impact (non-)disclosure decisions (e.g., Dechow and Sloan [1996]; Berger and Hann [2007]; Hoopes et al. [2018]; Bao et al. [2022, 2023]; Shi et al. [2023]; Jiang [2024]): forced labor scrutiny costs in global supply chains. Because maritime shipment records convey real-time and granular information on internationally traded goods for both publicly and privately held firms—essential for understanding collective corporate behavior in the context of social externalities—our setting provides a natural laboratory to study how forced labor scrutiny costs shape non-disclosure behavior. Our evidence that redactions are more likely to occur when forced labor risks are high is important in light of the growing demand for supply chain transparency as a means to mitigate firms' social externalities along their supply chains (Christensen et al. [2021]).

Second, we contribute to the emerging literature on supply chain transparency and corporate responsibility, with a special focus on societal issues. As governments and other external stakeholders increasingly demand greater information about the origins of firms' products, supply chain transparency and due diligence have become critical for firm operations (Sarfaty [2015]). In response to this demand, new US legislation has targeted human rights risks in supply chains, including Section 1502 of the Dodd-Frank Act and the California Transparency in Supply Chains Act. Prior research has focused on the impact of such non-financial (disclosure) regulations on various firm-level outcomes (e.g., Christensen et al. [2017]; Chen et al. [2018]; Rauter [2020]; She [2022]). In this paper, we focus on variations in firms' supply chain scrutiny costs related to forced labor risks and find that increases in such costs are associated with increased *non-disclosure* of international supply chain linkages. Importantly, this paper complements early policymaker evaluations of the trade enforcement impacts stemming from key developments in forced labor regulation such as the UFLPA (Greenfield et al. [2025]).

Third, we add to recent research investigating global supply chains using maritime shipment data (e.g., Jain et al. [2014]; Flaaen et al. [2023]; Bisetti et al. [2024]; Carter et al. [2024]; Ganapati et al. [2024]). Flaaen et al. [2023] highlight that firms' redactions can pose a limitation when using maritime shipment data to study the dynamics of trade relationships in certain research contexts.[9] In this paper, we leverage variation in transaction-level redactions to provide, to our knowledge, the first systematic evidence on the characteristics of redacted shipments and the underlying motivations behind identity redactions. Our evidence highlights that redactions in US shipment records are both material and non-random, which can impact the inferences external stakeholders draw relying on such data. For example, one may mistakenly conclude that firms have exited certain supplier relationships or regions when, in fact, these supply chain linkages have merely been concealed. Moreover, even without a causal interpretation of our results (and even if importers act in good faith), our investigation has an important policy implication: Because we document positive correlations between forced labor risks and redaction behavior, allowing firms to redact identities from maritime shipment records may inadvertently undermine the critical monitoring of supply chains by external stakeholders. Although firms provide (limited) voluntary disclosure of major supplier lists and certain supply chain disclosure mandates exist, granular manifest data on trade activity remains an important input to the external monitoring of social externalities resulting from global trade activity.

[9] Census data on international trade can be a suitable alternative data source, as census data is not limited to maritime transactions and also not subject to firm-level redaction decisions (Flaaen et al. [2023]). However, publicly available census data on international trade is aggregated, and more granular data is generally confidential with highly restricted access. As a result, maritime shipment data from CBP remains a uniquely granular data source to study the economic forces that determine trade relationships and the transparency thereof.

2. Institutional background

2.1. CBP's vessel manifest confidentiality program

For maritime shipments, the vessel manifest (also referred to as a cargo document) contains key shipment information. Particularly, the manifest includes information from the bills of lading for every order on board, such as the trade names of the supplier and importer, and the quantity and types of goods shipped. The manifest is mainly used by CBP to verify the cargo on board. Because the public can request and obtain granular information from these manifests via the Freedom of Information Act, CBP allows an importer to redact firm-identifying information such as the trade names and addresses of the importer and, if requested, the trade names and addresses of all suppliers before these records become publicly available. Various supply chain data providers, such as Panjiva or ImportGenius, regularly make use of their rights to systematically collect, process, and sell publicly available shipment records.[10]

To request manifest confidentiality, an importer can either send a letter or email to CBP or fill out an online application (i.e., the Vessel Manifest Confidentiality Online Application).[11] Any request is free of charge and granted for a forward-looking period of two years (importers cannot redact historical shipment records). Shipment records remain redacted indefinitely, and after the two-year period, importers wishing to continue in the manifest confidentiality program need to renew their request. Importantly, for CBP to grant confidentiality, importers must list any name variations in their request that would match the importer's trade name as written on a bill of lading.[12] To further conceal the identifying information of an importer's *suppliers*, the importer can simply request that all suppliers linked to its various importer name variations be hidden. Therefore, the importer does not

[10] For more information on CBP's vessel manifest confidentiality program see www.cbp.gov.

[11] Figure IA.1 provides screenshots of the online application. Online applications have been available since May 2020.

[12] Note that one importer (e.g., Walmart Inc.) can have many trade names (e.g., Wal Mart Stores Inc. or Wal-Mart PR) under which it imports goods. Because of incomplete name lists submitted by the redacting importer, misspellings in names on the bill of lading, or delayed renewal requests by the redacting importer (in addition to a backlog of requests prior to the online application system), redactions can be incomplete.

need to provide the exact supplier names to ensure confidentiality thereof. Upon the arrival of each shipment into a US port, redactions of identifying information are executed through exact name matches before making the remainder of shipment data publicly available.

Anecdotally, an oft-used argument in support of the program is that manifest confidentiality allows importers to conceal their trade relationships, thereby protecting their business strategies from competitors (Kienzle [2022]). For example, suppose an importer found a new supplier with excellent pricing and quality standards; that importer would not want to reveal its new supplier linkage and, by extension, details on granular shipment characteristics to its competitors. In support of the value of this data to external stakeholders to assess firm activities and competitive positions, Panjiva was recently acquired by S&P Global Market Intelligence.[13]

Importantly, however, importers do *not* need to provide evidence of competitive harm to apply for and be granted manifest confidentiality. As a consequence, redacted identities in shipment data could arise due to other supply chain opacity incentives, such as avoiding the costs associated with revealing trade linkages to regions, sectors, or suppliers that are suspected of using forced labor. In fact, public maritime shipment records have been used by external stakeholders to uncover questionable and possibly illicit supply chain relationships, thereby aiding regulatory enforcement.[14]

2.2. Forced labor in supply chains

US customs law banned firms from importing manufactured goods produced by certain categories of forced labor as early as the end of the nineteenth century and later codified this standard for all products (with some exceptions based on domestic demand) in the US Tariff Act of 1930 (CSR [2023]). Still, according to ILO [2024], 27.6 million people globally were in forced labor at any given time in 2021, suggesting an increase compared to their 2016

[13] See here for S&P's description of Panjiva as a tool to "identify potential business partners, keep track of competitors, or find investment signals."

[14] See here for a detailed example of the use of shipment data by activists to link post-sanction shipment activities to specific importers.

estimates. As a result, international efforts to combat forced labor in supply chains have become more urgent in recent years (Lobdell [2020]; Casey et al. [2023]). This urgency is fueled by demands from regulators, investors, and the general public for importers to adopt more environmentally and socially responsible practices that encompass the responsible sourcing of goods from suppliers.

Due to increased scrutiny, addressing forced labor has become a key focus of firms. To mitigate forced labor as a supply chain risk, importers often hire auditors to ensure that their suppliers comply with human rights and labor laws. However, these efforts can be inadequate due to auditors' time constraints in verifying labor standards and practices across all production sites (Rajagopalan [2024]). In addition, suppliers can conceal labor exploitation across multiple facilities or report unreliable and unverifiable information to importers. Because major importers may manage hundreds or even thousands of suppliers, the complexity of their sourcing network leads to significant costs associated with effective due diligence procedures aimed at addressing forced labor (Lobdell [2020]). Thus, even if importers act in good faith, statistics on forced labor suggest that their efforts are insufficient to effectively combat the issue (ILO [2024]).

Therefore, we conjecture that importers could prefer less transparency regarding their trade relationships if the revelation of forced labor risks in their supply chains is costly. For example, importers can incur costs related to the media, think tanks, and NGOs "naming and shaming" importers for unknowingly or knowingly being tied to forced labor conditions at supplier locations (Irwin et al. [2020]). Besides introducing reputational costs, such campaigns can further affect public sentiment leading to boycotts and follow-on regulatory pressure (Grant and Langpap [2024]).

2.3. International forced labor allegations

Although many goods and regions worldwide have been implicated in forced labor allegations, Xinjiang's cotton industry has faced particularly severe and ongoing international accusations of human rights abuses in recent years. Since 2014, Xinjiang's cotton industry has grown significantly: Xinjiang was responsible for over 90% of China's cotton production in 2022 (Weijia [2023]) and provides about 20% of the world's cotton (Uluyol [2023]). This growth in supply-chain importance of the region has been accompanied by international allegations of forced labor (e.g., Schmitz [2018]; Buckley and Ramzy [2019]; Leibold [2019]; Zenz [2019]; Congressional-Executive Commission on China [2021]).[15] In particular, there are claims of involuntary mass transfers of Uyghurs and other ethnic minority groups to factories between 2017 and 2019.[16]

In March 2020, the ASPI reported on these forced labor conditions, exposing global importers' linkages to Uyghur forced labor in suppliers' factories (ASPI [2020]). The report named more than 80 major companies including Abercrombie & Fitch, Gap, H&M, Ralph Lauren, and UNIQLO that could be directly or indirectly linked to suppliers engaged in forced labor practices and urged these importers to take immediate action to address human rights due diligence in their supply chain. Although many importers were quick to respond to these allegations and conducted independent investigations within their supply chain network, some importers remained silent.[17]

Importantly, although the report named only a subset of importers, it triggered a series of events with broad implications for various importers. Following the ASPI report, in July 2020, the US government issued a business advisory to importers with potential trade ties to Xinjiang, urging them to conduct rigorous due diligence to ensure their suppliers are compliant with human rights and labor laws (DOL [2020]; DOS [2020]). In the following months from August to December 2020, CBP issued Withhold Release Orders on goods from specific suppliers suspected of participating in state-sponsored forced labor (DHS [2020]).

[15] There have been global reactions to the alleged forced labor conditions in China but we focus our discussion on the US responses in line with our empirical analyses using data on US imports. Therefore, while the topic of forced labor allegations in Xinjiang is politically complex (e.g., Xinhua [2020]; Farge [2022]), the setting allows us to identify events that increase scrutiny on US importers with exposures to certain Chinese suppliers and industries suspected of forced labor.

[16] We collectively refer to Uyghurs but recognize that other members of ethnic minority groups are affected, such as Turkic-speaking Muslim minorities including the Kazakhs, Uzbeks, and Hui (see ASPI [2020]).

[17] See here for a collection of responses.

Then, in January 2021, CBP banned all cotton (and tomato) products that could be traced back to forced labor and the Xinjiang region (CBP [2021]).

To further strengthen existing laws—particularly the US Tariff Act of 1930—the US government signed the UFLPA in December 2021. The UFLPA came into effect in June 2022. The Act establishes a rebuttable presumption, indicating that the US government assumes goods produced (in whole or part) from Xinjiang or certain entities listed on the UFLPA entity list are made with forced labor and therefore prohibited unless the importer can clearly and convincingly prove otherwise (Swanson et al. [2022]; Uluyol [2023]). That is, if CBP has taken action against a certain shipment and the importer believes that the UFLPA does not apply to that shipment, the importer must provide convincing evidence that the shipment's goods cannot be traced back to Xinjiang or a supplier on the UFLPA entity list following CBP's guidelines.[18] As of June 2024, CBP has denied entry to 3,596 shipments.[19]

The series of events described above—starting with pressure from a non-partisan think tank (ASPI) and culminating in stringent regulatory actions and compliance measures in the US (UFLPA)—offers a unique setting to examine whether supply chain scrutiny of social externalities relates to redacted identities in shipment records. We expect that importers perceive redactions as a tool to lower the expected costs associated with third-party (and follow-on regulatory) pressures. Consequently, we predict that, after the onset of these events, the probability of redacted identities in maritime shipments increases.

In support of the relevance of these events in plausibly shaping redactions in maritime shipment data, it is notable that the Commercial Customs Operations Advisory Committee (COAC)—which is comprised of several major US importers—recently proposed that CBP should keep maritime shipment data *entirely* confidential (Goodman [2022]).[20] However, after strong opposition from several NGOs and legislators, the COAC withdrew their posi-

[18] See the UFLPA entity list and the UFLPA operational guidance for importers.
[19] See the UFLPA statistics of inspected shipments.
[20] See www.cbp.gov for more information on COAC.

tion. Importantly, external stakeholders emphasized that manifest confidentiality impedes the critical external monitoring of human rights abuses in supply chains and noted that firms' calls for increased data opacity are an attempt to undermine the efficacy of the UFLPA.[21] Highlighting the importance of external stakeholders' information gathering in shaping regulatory actions, legislators have advocated for increased public availability of manifest data to aid legal enforcement,[22] and CBP stresses the importance of tips in informing their forced labor investigations and audits, as stated on its website (Koscak [2022]):

> *CBP combats forced labor mainly through tips from the public, stakeholders, and other agencies since goods made by forced labor look the same as legitimate shipments.*

3. Data and sample selection

We obtain a data set of US maritime import transactions from S&P Global's Panjiva Supply Chain Intelligence. Panjiva collects US data directly from CBP through the Freedom of Information Act. Each transaction reflects a bill of lading which is a document typically issued by a transport company (i.e., carrier) to a shipper (i.e., supplier) certifying the receipt of the goods for transport from a departure location to a destination location and a consignee (i.e., importer). For each transaction (i.e., shipment), the following main elements are reported on the bill of lading for US imported goods: name and address of the importer, name and address of the supplier, description of the imported goods, weight of the imported goods, shipment arrival date, ports of lading and unlading, shipment origin, and shipment destination. Along with cleaning and partially standardizing the above fields, Panjiva utilizes other data and product descriptions to impute the following variables for

[21] See the open letter by NGOs here.

[22] See here the proposal by US Senators Sheldon Whitehouse (D-RI) and Bill Cassidy (R-LA) to make aircraft, rail, and truck trade data publicly available (alongside maritime shipment manifest data) to further enable the monitoring of supply chain activity and externalities. Generally, attempts to make all trade data publicly available have thus far failed (see, e.g., Panjiva, Inc. v. United States Customs and Border Protection, No. 19-118 (Second Circuit 2020)).

each transaction: the value in USD and the Harmonized System (HS) codes.[23]

Using US import data from 2013 through 2023, we start with a sample of about 133.3 million observations at the shipment level. For our analyses, we drop observations with missing HS codes, value, weight, country of shipment origin, region of port of unlading, and region of shipment destination.[24] We also eliminate observations when the two-digit HS codes are either 98 or 99 as those reflect special transactions and cannot be assigned a unique good type. To correct country fields for misspellings and non-English translations, we download the International Organization for Standardization's (ISO) country codes and names and manually construct a country dictionary. We then drop observations when the origin country remains unidentified.[25] Finally, by comparing Panjiva's parsed firm-identifying information fields with the original format of identifying information, we remove observations where importer or shipper identifying information is missing due to Panjiva's parsing limitations rather than redaction decisions.[26] This yields a final sample of over 104 million observations. The most frequently missing variables that remain in the data set are due to redactions from US importers reflected in missing firm-identifying information for either the importer or supplier, or both (given our research question, we keep these observations in the sample). We provide variable definitions in Appendix A.

[23] Panjiva uses a proprietary algorithm to translate product descriptions to HS codes. Because there can be multiple products in each bill of lading, we use the first imputed HS code listed by Panjiva. We retain six-digit HS codes that characterize a specific good (e.g., 85%-cotton yarn (520710)). On that basis, we construct two-digit HS codes that more generally characterize a type of good (e.g., cotton (52)).

[24] We drop observations with a missing shipment destination region variable, as these shipments are trans-shipped to the US but have a final destination elsewhere in the world (about 12 million observations). We further eliminate observations when the weight in metric tons on the bill of lading is rounded to zero (about 2.5 million observations). These observations are unlikely to reflect firm-level import transactions.

[25] In our country dictionary, we manually classify countries as unidentifiable when country fields contain two-letter abbreviations, unknown locations, and non-country labels such as "Other country", "The world", or "Countries not declared" (about 2,000 observations).

[26] On the other hand, Panjiva is sometimes able to recover redacted identifying information from other non-redacted fields (e.g., shipment notes). Such instances imply that redaction rates calculated in our data based on missing identifying information may reflect a lower bound of the prevalence of redaction decisions.

4. Descriptive statistics and prevalence of redacted identities

We first provide various descriptive statistics related to redacted identities to understand the prevalence of importers' redaction decisions at the shipment level. In Table 1, we report the proportion of shipments by year that have non-redacted and redacted identities. Overall, nearly 16% of shipments have redacted importer identities. This corresponds to 13% of the aggregate dollar value of maritime shipments (untabulated). Furthermore, this share has grown over time from 12% in 2013 to over 18% in 2023. Of these shipments with redacted importer identities, over 93% of them also have redacted supplier identities (untabulated). This is unsurprising given the relative ease with which importers can redact all supplier identities automatically upon request (see Figure IA.1).[27]

[Table 1]

In Table 2, we assess the share of shipments with redacted identities based on shipment origin region (Panel A) and US destination region (Panel B). We find that redacted identities are prevalent across all geographic regions. Notably, shipments from Eastern Asia, Micronesia, Northern Africa, South-Eastern Asia, and Western Asia have redaction rates above the overall sample mean of 16%. Although some of these regions represent a generally small number of shipments in the overall sample (Panel A, Column 2 reports the share of all observations that come from each region), note that over half of shipments (56.59%) come from Eastern Asia, and South-Eastern Asia represents the second largest origin region (10.27%). In addition, Panel B shows that redacted shipments have destinations across all regions of the US, most frequently in the New York and Pacific regions.

[Table 2]

[27] Although the overall redaction rate of 16% may seem relatively low given the ease with which redaction requests can be made, note that many of the remaining US shipments that are not redacted are received by shipment intermediaries (e.g., supply chain logistics companies). Specifically, we estimate that about 18% of shipments arrive into the US via an intermediary (untabulated), and the presence of such intermediaries naturally obfuscates the identifying information of the ultimate importer. Because the use of a shipment intermediary is not strictly a non-disclosure decision—it could simply reflect the operational need for intermediation in complex supply chains—we focus our analysis on the manifest confidentiality program.

In Table 3, we examine the distributional characteristics of two continuous variables in our data: the value of goods shipped (in US dollars, as imputed by Panjiva based on average per-unit-time prices) and the weight (in metric tons) of the goods shipped. In Panel A, we provide these summary statistics for the full sample of shipments. We find that the average shipment is worth about $167,000 and weighs 56 metric tons. However, there is significant skewness in the distribution of these variables; the median shipment is valued at $37,800 and weighs about nine metric tons. In Panels B and C, we split these summary statistics by shipments with non-redacted (Panel B) and redacted (Panel C) identities. We find that shipments with redacted identities tend to be of lesser value and weight, both in terms of the means and key percentiles of their distributions. Given the high skewness in the distributions, we use the quintiles of value and weight (Panel D) when utilizing them as control variables in our analyses.

[Table 3]

Overall, these descriptive statistics characterize the general prevalence of shipments with redacted identities. The share of shipments with redacted identities has grown over time, and nearly all of these redacted shipments have both importer and supplier identities redacted. Importantly, although various geographies exhibit shipments with redacted identities, these shares are not equally distributed. Next, we utilize other observable characteristics of shipments—such as the types of goods received, the date of arrival, and the country of origin—to examine the forces that incentivize importers to redact maritime shipment data.

5. Supply chain opacity incentives and redacted identities

There could be various incentives for non-disclosure that drive importers to redact their own identities in shipment records. We first investigate whether proxies of the presence of competitor-based proprietary costs—an oft-investigated force in driving non-disclosure decisions in other settings—relate to redacted shipments. Then, given CBP's longstanding stance against forced labor coupled with the continuing presence of forced labor risks in supply chains, we test whether shipments with higher risks of forced labor use (and thus subject to greater scrutiny from third parties) are more likely to have redacted identities.[28]

5.1. Competitor-based proprietary costs

5.1.1. Shipments of goods more exclusively related to the manufacturing sector

We test whether competitor-based proprietary costs correlate with redacted identities in maritime shipment data. According to survey data, the manufacturing sector especially relies on opacity as a key tool to protect proprietary information from competitors (Shackelford and Kindlon [2021]). Therefore, we test whether shipments of goods more exclusively related to the manufacturing sector (i.e., chemicals, plastics, technology) are more likely to have redacted identities relative to shipments of non-durable goods (e.g., footwear, toys).[29] We borrow a classification of two-digit HS codes into manufacturing and non-durable good categories (among other categories) from Liu et al. [2023] and estimate regressions of the following form at the shipment level (standard errors clustered at the destination region-good type level):[30]

$$\mathbb{1}(IdentityRedacted)_{i,t} = \beta \times \mathbb{1}(ManufacturingGood)_g + \delta_{or,dr,t} + \gamma_{v,w} + \epsilon_{i,t}\,. \tag{1}$$

Above, the subscript i reflects a class of shipments with the following characteristics: they

[28] In an untabulated analysis, we attempt to conduct *importer-level* determinant tests by imputing the timing of redaction decisions. We do so by (i) fuzzy matching Compustat names to Panjiva importer trade names, (ii) aggregating shipments at the importer-time level, and (iii) flagging sudden and persistent decreases in shipment quantities to very low (or zero) levels based on various distributional characteristics. We then regress an indicator variable for these imputed redaction periods on various determinants. In line with our shipment-level analysis, we find some evidence that redaction decisions are positively associated with proprietary costs and forced labor exposure proxies. On the other hand, they are negatively associated with information demand proxies such as institutional ownership (Cohen et al. [2023]). An important limitation of this analysis is that such imputation strategies may not reliably detect redaction periods for small importers, importers with various complex trade names, and importers that always redact.

[29] We exclude shipments of durable goods as these can sometimes relate to the manufacturing or retail sectors. Note that shipments of durable goods also exhibit higher redaction rates than non-durable goods (untabulated).

[30] An ideal clustering level for standard errors would nest the importer level, as importers make redaction decisions. However, because importer identities are inherently unobservable for redacted shipments, we cluster at a level that attempts to approximate importer location and industry.

originate from country o in region or of the world, have a final destination in region dr of the US, contain goods of type g based on two-digit HS code, and have value in quintile v and weight in quintile w. Because we run regressions at the shipment level, note that there can be multiple shipments within class i in each year-quarter t.

In examining whether shipments of manufacturing goods ($\mathbb{1}(ManufacturingGood)$) are more likely to have redacted identities ($\mathbb{1}(IdentityRedacted)$), we include two sets of fixed effects to help ensure that we compare the redaction rates of shipments related to importers that are more likely to be similar in their observable and unobservable characteristics.[31] $\delta_{or,dr,t}$ controls for time-varying supply chain shocks between geographical regions. $\gamma_{v,w}$ is a set of fixed effects that controls for the interaction of the value quintile and the weight quintile of the shipment, which relate to the insights in Table 3 that redacted shipments tend to be smaller. These fixed effects serve as a transaction-level size control. In this research design, β represents the variance-weighted average difference in identity redaction probabilities when comparing shipments of manufacturing goods versus non-durable goods coming from the same region in the world and arriving in the same destination region of the US at the same time (while also controlling for the shipment value and weight).

Results are in Table 4. Panel A provides summary statistics. In Panel B, Columns 1 and 2, we introduce fixed effects in a stepwise fashion, and in Column 3, we estimate the fully-specified design presented in Equation 1. We find that shipments of goods more exclusively related to the manufacturing sector have an incrementally higher redaction probability of 1.9%, relative to shipments of non-durable goods. These results are consistent with the notion of competitor-based proprietary costs driving importers' identity redactions.

[31] Because our tests are at the shipment level, we cannot ensure that our independent variables of interest perfectly classify different importers into separate levels (this would be ideal given that redaction is an importer-level decision). However, this classification concern should work against us finding meaningful differences between groups based on shipment-level characteristics.

5.1.2. Shipments reflecting new supply chain linkages

Instead of exploiting cross-sectional differences in the types of goods shipped, we next exploit the time series of shipments within a supply chain linkage to assess whether redacted identities are more likely to be present when a supply chain linkage is new. We argue that new supply chain linkages carry higher competitor-based proprietary costs because new supply chain linkages are less likely to be known otherwise by competitors (Kienzle [2022]). Specifically, we define a supply chain linkage as an *o-dr-g* (i.e., origin country-US destination region-good type) triplet. Within this linkage, we posit that the earliest set of shipments is more likely to carry competitor-based proprietary costs and therefore induce importers to redact identities from their shipments. To ensure that we do not errantly classify supply chain linkages as new due to our sample period starting in 2013, we only use shipment-level observations from 2016 onwards in our regression estimation of the following form (standard errors clustered at the destination region-good type level):

$$\mathbb{1}(IdentityRedacted)_{i,t} = \beta \times \mathbb{1}(NewLinkage)_{o,dr,g,t} + \iota_{o,dr,g} + \zeta_{g,t} + \gamma_{v,w} + \epsilon_{i,t}\,. \tag{2}$$

Above, $\mathbb{1}(NewLinkage)$ is defined as one when a shipment is within the first 500 (or the first 1,400 shipments) of a supply chain linkage, and as zero for any subsequent shipments within that same supply chain linkage.[32] The research design, which includes supply chain linkage ($\iota_{o,dr,g}$) and good type-time ($\zeta_{g,t}$) fixed effects, is similar to a (staggered) difference-in-differences design but in reverse. Many supply chain linkages have existed and traded heavily before 2016, and are therefore never treated.[33] Other supply chain linkages are treated for their earliest transactions that appear from 2016 onwards and untreated thereafter, with treatment initiation occurring at different points in time depending on the timing of the

[32]The cutoffs of 500 and 1,400 shipments roughly correspond to the 5th and 10th percentile of the overall timing-rank distribution of shipments (i.e., the 5th-(10th-)percentile shipment is approximately the 500th (1,400th) shipment in a supply chain linkage). In our tests, we drop any supply chain linkages that have fewer than those number of shipments such that the 500th and 1,400th shipments correspond to lower percentiles in the estimation sample.

[33]This large never-treated group reduces concerns associated with the negative weights assigned to counter-intuitive "early-to-late" comparisons in our staggered difference-in-differences design (as these weights will be relatively small).

very first shipment. In this research design, β represents the variance-weighted average difference in the redaction probability for new shipments relative to later shipments in the same linkage, as compared to changes within control linkages for the same type of good (but traveling from and to different locations) over that time period (while also controlling for the shipment value and weight using $\gamma_{v,w}$).

Results are in Table 4, Panel C. In Columns 2 and 4, we estimate the fully-specified design presented in Equation 2 for cutoffs of 500 shipments and 1,400 shipments, respectively. In both cases, we find that new shipments within a supply chain linkage are more likely to have redacted identities. This effect is generally stronger when the cutoff is 500 shipments than when the cutoff is 1,400 shipments, implying that the very earliest shipments in a supply chain linkage carry the highest competitor-based proprietary costs.

[Table 4]

Overall, the results of Table 4 support the notion that competitor-based proprietary costs are a factor in driving importers' redaction decisions in shipment records, consistent with the literature on redactions in other settings (e.g., Verrecchia and Weber [2006]; Li et al. [2018]; Glaeser [2018]; Chen et al. [2022]). In our next set of tests, we assess whether proxies of other supply chain scrutiny costs—those associated with the revelation of forced labor risks—relate to redacted identities in maritime shipment data.

5.2. Supply chain scrutiny costs associated with forced labor risks

To assess whether supply chain scrutiny related to the revelation of forced labor risks is costly to importers and therefore incentivizes redaction decisions, we rely on third-party measures of forced labor risks in shipment origin countries by Walk Free. Walk Free is an international human rights group formed in 2010. They calculate scores of country-level forced labor vulnerability and government responses to forced labor risks in their Global Slavery Index. Measures of vulnerability are based on factors such as governance issues, inequality, and disenfranchised groups; measures of government responses are based on the treatment of survivors of modern slavery, criminal justice mechanisms, and institutional strength.[34] We create indicator variables based on these scores and estimate regressions of the following form at the shipment level (standard errors clustered at the destination region-good type level):

$$\mathbb{1}(IdentityRedacted)_{i,t} = \beta \times \mathbb{1}(ForcedLaborIndicator)_o + \phi_{dr,g,t} + \gamma_{v,w} + \epsilon_{i,t}\,. \tag{3}$$

$\mathbb{1}(ForcedLaborIndicator)$ can be one of two measures. $\mathbb{1}(HighVulnerabilityCountry)$ is based on the vulnerability score from Walk Free. This variable is set to one for shipments from countries where the vulnerability score is greater than or equal to 30, and set to zero otherwise. $\mathbb{1}(LowGovernmentResponseCountry)$ is set to one for shipments from countries where the government response score is less than or equal to 50, and set to zero otherwise.[35] Because shipment origin is the key source of variation in this test, we apply a different set of control variables compared to prior tests. Namely, the set of fixed effects $\phi_{dr,g,t}$ forms the primary cross-section of comparison, that is, goods of the same type arriving in the same year-quarter to the same destination region of the US. In this research design, β represents the variance-weighted average difference in the redaction probability between high-forced-labor-risk and low-forced-labor-risk shipments within these cross-sections (while also controlling for shipment value and weight).

Results are in Table 5. Panel A provides summary statistics for explanatory variables. In Panel B, the explanatory variable is based on the vulnerability score, and in Column 2, we estimate the design in Equation 3. We find that shipments from origin countries with high forced labor vulnerabilities are more likely to have redacted identities. In Panel C, the explanatory variable is based on the government response score. Shipments from countries

[34] See www.walkfree.org for details on Walk Free and their data methodology.

[35] The cutoff of 30 for $\mathbb{1}(HighVulnerabilityCountry)$ reflects a natural break in the distribution where there is a large gap in scores between adjacently ranked countries. Unlike the distribution of vulnerability scores, the government response score has no natural break in the distribution, leading to a choice of 50 (the midpoint of the score range). When a country score is missing, the relevant indicator variable is set to missing. Note that these indicator variables have overlapping values for over 75% of the countries, indicating a strong relation between forced labor vulnerability and limited government response thereto.

whose governments have limited responses to forced labor risks are also more likely to be redacted.[36]

[Table 5]

Overall, the tests in Table 5 provide evidence consistent with the notion that supply chain scrutiny related to social externalities could incentivize importers exposed to forced labor risks to obfuscate their trade relationships. In our final set of tests, we focus on a series of events that materially increased supply chain scrutiny as it relates to forced labor risks in cotton and apparel production in China. This "shock," evaluated within a difference-in-difference-in-differences design, allows us to draw insights about how changes in supply chain scrutiny costs relate to changes in redacted identities.

6. International forced labor allegations and redacted identities

As outlined in Section 2.3, a series of supply chain scrutiny events—starting with public scrutiny from a non-partisan think tank and culminating with continued regulatory scrutiny—recently affected the attention paid to human rights due diligence in cotton and apparel production in China. We study whether these events (starting after March 2020) affected the probability of redacted identities for imports of cotton- and apparel-related goods from China.

6.1. Research design

To assess how changes in supply chain scrutiny relate to changes in identity redaction probabilities, we estimate a difference-in-difference-in-differences design. The primary change we are interested in relates to the redaction of imports of cotton- and apparel-related goods

[36] Note that our explanatory measures of forced labor risks vary at the origin-country level. Because our analysis is at the shipment level, the estimation generally places higher weight on shipments that originate from key origin countries (such as China). In an untabulated analysis, we collapse the data to the origin country-year quarter level, thereby assigning equal weight to each of these country-year-quarter cells. We similarly find a strong relation between forced labor risks and redacted shipments.

from China around these events. However, concurrent events such as the COVID-19 pandemic (Liu et al. [2023]) and the "trade war" between the US and China (Fajgelbaum et al. [2024]) could have impacts on our outcome variable beyond the forces related to our studied scrutiny events. For example, disruptions caused by COVID-19 or the "trade war" could generally lead more resilient importers to continue trading with China across all types of imported goods, and resilient importers may generally choose redaction. Therefore, to better isolate the impact of scrutiny related to forced labor risks in the cotton and apparel supply chains, we estimate regressions of the following form from 2018 to 2023 (standard errors clustered at the destination region-good type level):

$$\begin{aligned}\mathbb{1}(IdentityRedacted)_{i,t} = \beta \times [\mathbb{1}(Cotton, Apparel)_a \times \mathbb{1}(FromChina)_c \times PostASPI_t] \\ + \alpha_{a,c} + \kappa_{a,t} + \varphi_{c,t} + \epsilon_{i,t}\,.\end{aligned} \tag{4}$$

The difference-in-difference-in-differences design above has four cells of interest, captured by the fixed effects $\alpha_{a,c}$: (i) cotton and apparel shipments from China ($a = 1$, $c = 1$), (ii) cotton and apparel shipments from elsewhere in the world ($a = 1$, $c = 0$), (iii) other shipments from China ($a = 0$, $c = 1$), and (iv) other shipments from elsewhere in the world ($a = 0$, $c = 0$). Based on the good types identified in Murphy [2021] (see their Annex C), cotton and apparel shipments are defined as any shipments with two-digit HS codes 52, 60, 61, and 62. These shipments are treated when they come from China (cell (i)), and they are control observations when they come from elsewhere in the world (cell (ii)). Because the same importers that source cotton and apparel goods from China may source other non-durable goods from China (and importer overlap across our cells represents a SUTVA concern), we define shipments of manufacturing and durable goods as our non-cotton/apparel control observations (i.e., cells (iii) and (iv)).[37] Next, $PostASPI$ is an indicator variable taking a value of one in 2020Q2 and onwards, marking the period after the ASPI publishes its report scrutinizing importers'

[37]Because suppliers of certain electronics goods were targeted in the ASPI report (but these goods were not related to all events in our analysis), we exclude shipments of mobile phones (four-digit HS codes 8517) and laptops and computers (four-digit HS codes 8471) from the estimation. Results are similar if we add combinations of these and other scrutinized goods that are relevant to at least one of the events in our analysis (e.g., automobiles and footwear) to the treatment group (untabulated).

linkages to forced labor risks. Along with the time-invariant fixed effects $\alpha_{a,c}$, we include two sets of time fixed effects—$\kappa_{a,t}$ for shipments of cotton and apparel (or other) goods from anywhere at time t and $\varphi_{c,t}$ for shipments of any goods originating from China (or elsewhere) at time t—that complete the difference-in-difference-in-differences design.

An important assumption underlying our design is a parallel trends assumption between difference-in-differences (DiD). Specifically, our design uses the DiD between (iii) and (iv) as the relative group over time for the DiD between (i) and (ii) (one can also interpret this design as the DiD between (ii) and (iv) being the relative group over time for the DiD between (i) and (iii)). In this design, β cannot be easily explained by overall differences in redactions between shipments from China and the rest of the world during this period or overall differences in redactions between cotton/apparel-related shipments and other shipments during this period. Conditional on the parallel trends assumption not being violated, the coefficient on β represents the variance-weighted average treatment effect on the treated (VWATT) of forced labor scrutiny on the probability of identity redactions in shipment records.[38]

6.2. Results

In Table 6, we estimate the effects of the scrutiny events related to forced labor allegations in China on redacted shipment identities.[39] In Column 1, we estimate a simple DiD design. In this design, the pooled control group contains shipments of manufacturing and durable goods from China and elsewhere in the world, as well as shipments of cotton/apparel goods from elsewhere in the world. The coefficient on $\mathbb{1}(Cotton, Apparel) \times \mathbb{1}(FromChina)$ is akin to a treatment indicator variable: cotton and apparel shipments from China have a higher, but statistically insignificant, redaction probability prior to April 2020. The coefficient on $PostASPI$ implies that, in and after April 2020, the redaction probability across

[38]SUTVA is another key assumption underlying our design that is likely to be violated to some degree because a single importer can receive shipments from many geographies and for many types of goods. Note that this version of a SUTVA concern would work against estimating a positive β. Moreover, we address such concerns by limiting our non-cotton/apparel control observations to manufacturing and durable goods.

[39]Table IA.1 provides a series of robustness tests related to the analysis in this section.

all control shipments in the sample is higher. Importantly, the coefficient of interest on $\mathbb{1}(Cotton, Apparel) \times \mathbb{1}(FromChina) \times PostASPI$ reveals an increase in the redaction probability for cotton and apparel shipments from China that is nearly three times larger than the increase observed in the pooled control group.

Although the specification of Column 1 provides some helpful insights on pre-period levels of redaction probabilities for treated shipments and post-period changes in the redaction probability of control shipments, it does not carefully address the possible endogeneity concerns outlined in Section 6.1. In Column 2, we therefore implement the design presented in Equation 4. This difference-in-difference-in-differences design estimates a β with greater statistical precision and a larger magnitude, relative to Column 1. After the onset of supply chain scrutiny events regarding forced labor risks in China, we find an increase in the probability of identity redaction for cotton/apparel shipments from China of about 5%. In Column 3, we introduce further granularity in our fixed effects by adding interactions with two-digit HS code (a more granular version of the fixed effects a) and destination region. Note that doing so nests the time-invariant fixed effects within our standard error clusters. We also add destination region-time fixed effects to control for any disruption shocks that may alter the nature of shipments that generally arrive at certain US destinations. The coefficient estimate remains similar. In Column 4, we introduce the value quintile-weight quintile control for shipment size, and we continue to estimate a similarly sized effect of about 5%.[40]

[Table 6]

Note that the above estimation pooled the entire post-period after the ASPI report into one estimate. However, as outlined in Section 2.3, there are several distinct events that

[40] In an untabulated analysis, we collapse the data to the From China (Y/N)-good type-year quarter level and assess whether the number of Chinese cotton and apparel shipments decreases after the ASPI report using a difference-in-difference-in-differences research design. We find a statistically insignificant decrease in the number of cotton and apparel shipments from China, and in event time, this decrease is more consistent after the UFLPA. This result aligns with the UFLPA policy analysis in Greenfield et al. [2025], indicating that, on average, importers may reduce the amount of scrutinized goods they source from China after regulatory scrutiny comes into play.

could impact the probability that shipments have redacted identities. To the extent that these events occur with sufficient time gaps, we can assess whether changes in the redaction probability differ over the time series in relation to the onset of each event. Figure 1 presents the results of the event-time estimation in which we modify the specification presented in Equation 4 and split $PostASPI$ into a series of pre- and post-periods. Given that the frequency of shipments can be sparse at an importer level (Flaaen et al. [2023]) and redaction decisions have a multi-period effect, we pool time periods either into two- or three-quarter segments, accounting for the timing of the ASPI report.[41]

We find that in the lead-up to the ASPI report, there is no clear time trend in the difference in the redaction probability between treated and control shipments, providing confidence in the quality of the control observations and the validity of our parallel trends assumption. In the three-quarter period after the ASPI report, we document an increase in cotton/apparel-related redaction probabilities from China (relative to control observations). This difference persists after CBP banned cotton (and tomato) imports from Xinjiang in 2021. Although the effect is attenuated in 2022, note that there is an increase in the redaction probability when the UFLPA comes into effect. This increase is statistically significant and persistent in 2023 (relative to pre-ASPI-report levels).

[Figure 1]

The results above are consistent with the notion that increases in supply chain scrutiny related to forced labor risks leads to higher redaction probabilities in shipments of scrutinized goods. However, because we conduct our analysis at the shipment level and do not know which (and when) importers have filed redaction requests, these results should be interpreted with caution.

Specifically, our results can be explained by two non-mutually exclusive channels. First, the increase in forced labor scrutiny in China could induce importers to file new redaction

[41] Note that effects in event time reflect the timing with which importers receive shipments (not necessarily the timing of redaction decisions). That is, the variation over time in the post-period reflects the composition of shipments to non-redacting versus redacting importers at that point in time.

requests while they continue to source cotton and apparel from China. As a result, the increase in the redaction probability over the post-period would reflect these intensive-margin adjustments in redaction requests. Second, the increase in forced labor scrutiny could have made it more difficult for non-redacting importers to continue sourcing cotton and apparel from China; therefore, the remaining importers in that market are more likely to have redacted identities. This extensive-margin effect would similarly show an average increase in the redaction probability over the post-period. Importantly, both channels convey a similar takeaway: because redacted shipments are likely to carry lower supply chain scrutiny costs, there is an increase in the probability of redacted shipments when such scrutiny costs increase.

In Internet Appendix Table IA.2, we provide additional evidence that supports the idea that the increase in redacted shipments after the forced labor scrutiny events can, at least in part, be attributed to new redaction requests. Specifically, we focus on a new outcome variable that captures the complete disappearance of importers' trade names in the post-ASPI period, leveraging Panjiva's unique trade name identifier to track trade names over time (note that this field is missing for redacted shipments). The intuition behind this test is as follows: if importers sourcing cotton and apparel goods from China begin redacting identities in the post-period due to heightened supply chain scrutiny, their associated trade names should *completely* disappear from all observable supply chains during this time. Although the complexity of importers' multiple trade names and the presence of privately held firms in our sample limit our ability to link trade names to specific importers, we use Panjiva data to classify each trade name's exposure to Chinese cotton and apparel goods in the pre-ASPI period and then examine whether exposed trade names completely disappear from shipment records in the post-ASPI period.

To perform this analysis, we first collapse the shipment-level data to the trade name level. We then classify a trade name as exposed if 50% or more of its pre-ASPI shipments consisted of cotton and apparel shipments originating from China. This indicator variable serves as our primary explanatory variable, distinguishing trade names that are exposed or not exposed to forced labor scrutiny in our setting. We define our outcome variable as an indicator variable set equal to one if the trade name completely disappears from Panjiva data in the post-ASPI period (i.e., there are no new shipment records associated with that trade name), and zero otherwise. Since trade name disappearance could indicate importers ceasing operations rather than making new redaction requests, we restrict our sample to trade names actively engaged in importing—specifically, those equal to or above the 90th percentile of the distribution of the number of shipments from 2018Q1 through 2020Q1 (i.e., the pre-ASPI period). For these trade names, we argue that their complete disappearance from all supply chains in the post-ASPI period is more likely due to new redaction decisions.

In Internet Appendix Table IA.2, Column 1 of Panel A, we find that non-exposed trade names exhibit a baseline disappearance rate of 5.3%, possibly reflecting the general rate of business closures following the COVID-19 pandemic. Importantly, however, exposed trade names are significantly more likely to disappear with an incremental disappearance rate of 8.5% compared to non-exposed trade names. In Columns 2 through 4, we introduce size quantile controls and further restrict the estimation sample to ensure counterfactual non-exposed firms are more similar to exposed firms in their sourcing characteristics, and in Panel B, we exchange the exposure indicator variable with a continuous exposure variable. We consistently find similar results. The collection of evidence implies that the recent events increasing scrutiny of importers' forced labor risks resulted in new redaction decisions, which contribute to the overall effects documented in Table 6.

7. Conclusion

CBP permits firms to request the redaction of their own and their suppliers' identifying information in transaction-level shipment records. Studying the characteristics of redacted identities in shipment records, we find that approximately 16% of shipments from 2013 to 2023 contain redacted identities, with substantial variation over time, across origin regions, and based on other shipment characteristics. Then, we not only connect to prior research that identifies proprietary cost concerns as an important motive of firms' (non-)disclosure decisions, but also provide evidence on a significant but understudied force: forced labor scrutiny costs. In doing so, we exploit a series of recent events that increased public and regulatory scrutiny of importers' human rights due diligence in supply chains. We document a notable increase in the probability of redacted identities after these events in the most scrutinized supply chains. Overall, our results have important policy implications: redacted identities in maritime shipment data could hinder the complete and critical monitoring of supply chain activities in the areas most exposed to social externalities like forced labor. Such insights are important as US policymakers continue to take material steps in combating forced labor in part through business stakeholder engagement and supply chain transparency initiatives.

We note the following caveats to our analyses and interpretations. First, as mentioned throughout the paper, our shipment-level analyses do not directly translate into importer-level redaction decisions. Therefore, results on redacted shipments throughout the paper should be interpreted as reflecting both (i) importers' decisions to redact and (ii) redacted importers' market share in particular supply chains. Second, we design our analyses to be able to speak to forced labor *risks* and the resulting pressures of the revelation thereof in importers' supply chains. We do not observe the actual occurrence of forced labor in suppliers' production facilities that can be linked to importers. Therefore, our results do *not* imply that importers choose to redact their identities to systematically and knowingly hide the presence of forced labor in supply chains. Third, the inferences from our difference-in-difference-in-differences analysis rely on the parallel trends assumption between difference-in-differences not being violated. Although an inspection of the pre-period trends provides some support for the validity of the parallel trends assumption, there could be differential effects of COVID-19 and trade relations specifically on cotton and apparel shipments unaccounted for by the research design. Fourth, we study redaction decisions in maritime shipment data and find that 16% of maritime shipments have redacted identities. However, maritime shipments make up only 45% of traded value in the US, with the remainder of shipments coming by air or land. Data on air, truck, and rail shipments are entirely confidential, meaning our results only speak to a portion (albeit a sizeable one) of US import activity.

References

ASPI (2020). Uyghurs for sale. *Australian Strategic Policy Institute*, (Policy Brief, Report No. 26).

Bao, D., Kim, Y., and Su, L. N. (2022). Do Firms Redact Information from Material Contracts to Conceal Bad News? *The Accounting Review*, 97(5):29–57.

Bao, D., Myers, L. A., and Su, L. N. (2023). Tax-Related Disclosure Costs: Evidence from Redactions in Material Contracts. *SSRN Electronic Journal*.

Baron, D. P. (2001). Private Politics, Corporate Social Responsibility, and Integrated Strategy. *Journal of Economics and Management Strategy*, 10(1):7–45.

Baron, D. P. (2003). Private Politics. *Journal of Economics*, 12(1):31–66.

Bateman, A. and Bonanni, L. (2019). What Supply Chain Transparency Really Means. *Harvard Business Review*.

Berger, P. G. and Hann, R. N. (2007). Segment Profitability and the Proprietary and Agency Costs of Disclosure. *The Accounting Review*, 82(4):869–906.

Bisetti, E., She, G., and Zaldokas, A. (2024). ESG Shocks in Global Supply Chains. *SSRN Electronic Journal*.

Buckley, C. and Ramzy, A. (2019). Inside China's Push to Turn Muslim Minorities Into an Army of Workers. *The New York Times*.

Carter, M. E., Lee, L. F., and Yu, E. (2024). Real Effects of the Proposed SEC Climate Disclosure Rule. *SSRN Electronic Journal*.

Casey, C. A., Cimino-Isaacs, C. D., and Weber, M. A. (2023). Section 307 and Imports Produced by Forced Labor. *Congressional Research Service Report*.

CBP (2021). CBP Issues Region-Wide Withhold Release Order on Products Made by Slave Labor in Xinjiang. *US Customs and Border Protection*.

Chen, G., Tian, X. S., and Yu, M. (2022). Redact to protect? Customers' incentive to protect information and suppliers' disclosure strategies. *Journal of Accounting and Economics*, 74(1):101490.

Chen, Y.-C., Hung, M., and Wang, Y. (2018). The effect of mandatory CSR disclosure on firm profitability and social externalities: Evidence from China. *Journal of Accounting and Economics*, 65(1):169–190.

Christensen, H., Hail, L., and Leuz, C. (2021). Mandatory CSR and sustainability reporting: Economic analysis and literature review. *Review of Accounting Studies*, 26:1–73.

Christensen, H. B., Floyd, E., Liu, L. Y., and Maffett, M. (2017). The real effects of mandated information on social responsibility in financial reports: Evidence from mine-safety records. *Journal of Accounting and Economics*, 64(2-3):284–304.

Cohen, S., Kadach, I., and Ormazabal, G. (2023). Institutional investors, climate disclosure, and carbon emissions. *Journal of Accounting and Economics*, 76(2-3):101640.

Congressional-Executive Commission on China (2021). Global Supply Chains, Forced Labor, and the Xinjiang Uyghur Autonomous Region.

CSR (2023). Section 307 and Imports Produced by Forced Labor. *Congressional Research Service*.

Dechow, P. M. and Sloan, R. G. (1996). Economic Consequences of Accounting for Stock-Based Compensation. *Journal of Accounting Research*, 34:1–20.

DHS (2020). DHS Cracks Down on Goods Produced by China's State-Sponsored Forced Labor. *US Department of Homeland Security*.

DOL (2020). Against Their Will: The Situation in Xinjiang. *US Department of Labor*.

DOS (2020). Xinjiang Supply Chain Business Advisory. *US Department of State*.

Egorovy, G. and Harstadz, B. (2017). Private Politics and Public Regulation. *The Review of Economic Studies*, 84:1652–1682.

Ellis, J. A., Fee, C. E., and Thomas, S. E. (2012). Proprietary costs and the disclosure of information about customers. *Journal of Accounting Research*, 50(3):685–727.

Fajgelbaum, P., Goldberg, P., Kennedy, P., Khandelwal, A., and Taglioni, D. (2024). The US-China trade war and global reallocations. *American Economic Review: Insights*, 6(2):295–312.

Farge, E. (2022). U.N. body rejects debate on China's treatment of Uyghur Muslims in blow to West. *Thomson Reuters*.

Flaaen, A., Haberkorn, F., Lewis, L., Monken, A., Pierce, J., Rhodes, R., and Yi, M. (2023). Bill of lading data in international trade research with an application to the COVID-19 pandemic. *Review of International Economics*, 31(3):1146–1172.

Ganapati, S., Wong, W. F., and Ziv, O. (2024). Entrepôt: Hubs, Scale, and Trade Costs. *American Economic Journal: Macroeconomics*, 16(4):239–278.

Glaeser, S. (2018). The effects of proprietary information on corporate disclosure and transparency: Evidence from trade secrets. *Journal of Accounting and Economics*, 66(1):163–193.

Goodman, J. (2022). US businesses propose hiding trade data used to trace abuse. *AP News*.

Grant, L. and Langpap, C. (2024). The Roles of Environmental Groups in Economics. *Review of Environmental Economics and Policy*, 18(2):000–000.

Greenfield, V. A., Sytsma, T., Kerrigan, A., Buenaventura, M., Patel, K. V., Steiner, M., Meredith, M., Scruggs, A., Hoak, L., Giglio, K., Hicks, D., and Welburn, J. W. (2025). *Forced Labor in Global Supply Chains: Trade Enforcement Impacts and Opportunities*. RAND Corporation, Santa Monica, CA.

Hoopes, J., Robinson, L., and Slemrod, J. (2018). Public tax-return disclosure. *Journal of Accounting and Economics*, 66(1):142–162.

ILO (2024). *Profits and poverty: The economics of forced labour*. International Labour Office, Geneva, second edition edition.

Irwin, P., Ilham, J., Cranston, C., and Sobhi Damasio, N. (2020). Press Release: 180+ Orgs Demand Apparel Brands End Complicity in Uyghur Forced Labour. *End Uyghur Forced Labour*.

Jain, N., Girotra, K., and Netessine, S. (2014). Managing Global Sourcing: Inventory Performance. *Management Science*, 60(5):1202–1222.

Jiang, Y. A. (2024). Mandatory Environmental Disclosure with Options to Withhold Trade Secrets. *SSRN Electronic Journal*.

Kienzle, N. (2022). Manifest Confidentiality: A Legal Way to Hide Import Records. *USA Customs Clearance*.

Koscak, P. (2022). CBP Takes Aim at Forced Labor. *US Customs and Border Protection*.

Leibold, J. (2019). Despite China's denials, its treatment of the Uyghurs should be called what it is: cultural genocide. *The Conversation*.

Li, Y., Lin, Y., and Zhang, L. (2018). Trade secrets law and corporate disclosure: Causal evidence on the proprietary cost hypothesis. *Journal of Accounting Research*, 56(1):265–308.

Liu, E., Smirnyagin, V., and Tsyvinski, A. (2023). Supply Disruptions Index (SDI): Data and Methodology. *SSRN Electronic Journal*.

Lobdell, K. (2020). Forced labor and supply chain risk: What companies need to know. *Thomson Reuters*.

Murphy, L. T. (2021). Laundering Cotton: How Xinjiang Cotton Is Obscured In International Supply Chains. *Sheffield Hallam University Helena Kennedy Center for Interna-*

tional Justice.

Rajagopalan, M. (2024). Serious labor abuses. *The New York Times*.

Rauter, T. (2020). The Effect of Mandatory Extraction Payment Disclosures on Corporate Payment and Investment Policies Abroad. *Journal of Accounting Research*, 58(5):1075–1116.

Sarfaty, G. A. (2015). Shining light on global supply chains. *Harvard International Law Journal*, 56(2):419–463.

Schmitz, R. (2018). Ex-Detainee Describes Torture In China's Xinjiang Re-Education Camp. *NPR*.

Shackelford, B. and Kindlon, A. (2021). Nondisclosure Agreements, Trade Secrets, and Trademarks Considered Very Important to More U.S. Businesses than Were Patents or Copyrights in 2017. *National Center for Science and Engineering Statitics*.

She, G. (2022). The real effects of mandatory nonfinancial disclosure: Evidence from supply chain transparency. *The Accounting Review*, 97(5):399–425.

Shi, Y., Wu, J., Zhang, Y., and Zhou, Y. (2023). Green Image Management in Supply Chains: Strategic Disclosure of Corporate Suppliers. *SSRN Electronic Journal*.

Swanson, A., Edmondson, C., and Wong, E. (2022). U.S. Effort to Combat Forced Labor Targets Corporate China Ties. *The New York Times*.

Uluyol, Y. (2023). Tailoring Responsibility: Tracing Apparel Supply Chains from the Uyghur Region to Europe. Technical report, Sheffield Hallam University Helena Kennedy Centre for International Justice, Sheffield, UK.

Verrecchia, R. E. and Weber, J. (2006). Redacted Disclosure. *Journal of Accounting Research*, 44(4):791–814.

Wagenhofer, A. (1990). Voluntary disclosure with a strategic opponent. *Journal of Accounting and Economics*, 12(4):341–363.

Weijia, H. (2023). Cotton trading center key step to establish textile clusters in Xinjiang. *Global Times*.

Xinhua (2020). China slams U.S. for blocking cotton imports from Xinjiang. *Xinhua Net*.

Zenz, A. (2019). Brainwashing, police guards and coercive internment: evidence from Chinese government documents about the nature and extent of Xinjiang's 'vocational training internment camps'. *Journal of Political Risk*, 7(7):1.

Appendix A. Variable definitions

Variable	Definition and measurement
Origin country	The country from which the shipment originated, parsed by Panjiva and manually cleaned. (Source: Panjiva)
Origin region	The geographical region of the origin country. (Source: UN Statistics Division)
Destination region	The geographical US region of the shipment's final destination. (Source: Panjiva)
Good type	The two-digit HS code for the shipped goods, reflecting broad product categories. Panjiva imputes these codes from textual descriptions using a proprietary algorithm. If multiple HS codes are associated with a single shipment, we use the first listed HS code. (Source: Panjiva)
Value of goods	The estimated dollar value of shipped goods. Panjiva imputes this value based on shipment size, good type/description, and average per-unit-time prices in US dollars. (Source: Panjiva)
Weight	The weight of shipped goods in metric tons. (Source: Panjiva)
Value quintile	A five-bucket sort of the value of shipped goods.
Weight quintile	A five-bucket sort of the weight of shipped goods.
Manufacturing goods	Goods with the following two-digit HS codes: 28-29, 31-41, 44-48, 54, 68, 73, 83-89, 96 (based on the classification in Liu et al. [2023]).
Durable goods	Goods with the following two-digit HS codes: 49, 57, 69-70, 82, 91-94, 97 (based on the classification in Liu et al. [2023]).
Non-durable goods	Goods with the following two-digit HS codes: 02, 04, 05, 07-10, 12-22, 24-25, 42-43, 50-53, 55-56, 58-67, 95 (based on the classification in Liu et al. [2023]).
$\mathbb{1}$(Identity Redacted)	An indicator variable equal to one if the importer's identifying information is redacted from a shipment (such as the trade name and address), and zero otherwise.
$\mathbb{1}$(Manufacturing Good)	An indicator variable equal to one if the shipped goods are classified as manufacturing goods, and zero otherwise.
$\mathbb{1}$(New Linkage)	An indicator variable equal to one if a shipment is one of the first 500 (or 1,400) shipments within a supply chain linkage (a linkage is defined as an origin country-destination region-good type triplet), and zero otherwise.
$\mathbb{1}$(High Vulnerability Country)	An indicator variable equal to one if a shipment's origin country has a forced labor vulnerability score greater than or equal to 30%, and zero otherwise. The cutoff of 30% reflects a natural break in the distribution of the vulnerability scores. (Source: Walk Free)
$\mathbb{1}$(Low Government Response Country)	An indicator variable equal to one if a shipment's origin country has a forced labor government response score less than or equal to 50%, and zero otherwise. Because there is no natural break in the distribution of the government-response score, the cutoff is set to the midpoint of the distribution. (Source: Walk Free)
$\mathbb{1}$(Cotton, Apparel)	An indicator variable equal to one if the shipped goods are classified as cotton or apparel, and zero otherwise. Goods with the following HS codes are cotton- or apparel-related: 52, 60, 61, or 62 (based on the classification in Murphy [2021]).
$\mathbb{1}$(From China)	An indicator variable equal to one if the shipment's origin country is China, and zero otherwise.
PostASPI	An indicator variable equal to one for shipments in and after April 2020, and zero otherwise.

Figure 1

Forced labor scrutiny and redacted identities over time

This figure modifies the specification in Table 6, Column 2 (Equation 4) into the following event time specification:

$$\mathbb{1}(IdentityRedacted)_{i,t} = \sum_{p=-4}^{6} \beta_p \times [\mathbb{1}(Cotton, Apparel)_a * \mathbb{1}(FromChina)_c * ASPI_p] + \alpha_{a,c} + \kappa_{a,t} + \varphi_{c,t} + \epsilon_{i,t} \ ,$$

where β_p reflects estimates of the difference between treatment and control observations for each of ten time periods relative to the difference in the period just before the impact of the ASPI forced labor report ($p = -1$). Above, the subscript i reflects a class of shipments with the following characteristics: they originate from country o in region or of the world, have a final destination in region dr of the US, contain goods of type g based on two-digit HS code, and have value in quintile v and weight in quintile w. a indicates whether a good type is (or is not) cotton/apparel, and c indicates whether the shipment of that good originates (or does not originate) from China. Key variables are defined in Appendix A. Standard errors are clustered at the destination region-good type level, and the standard error bars indicate a 90% confidence interval.

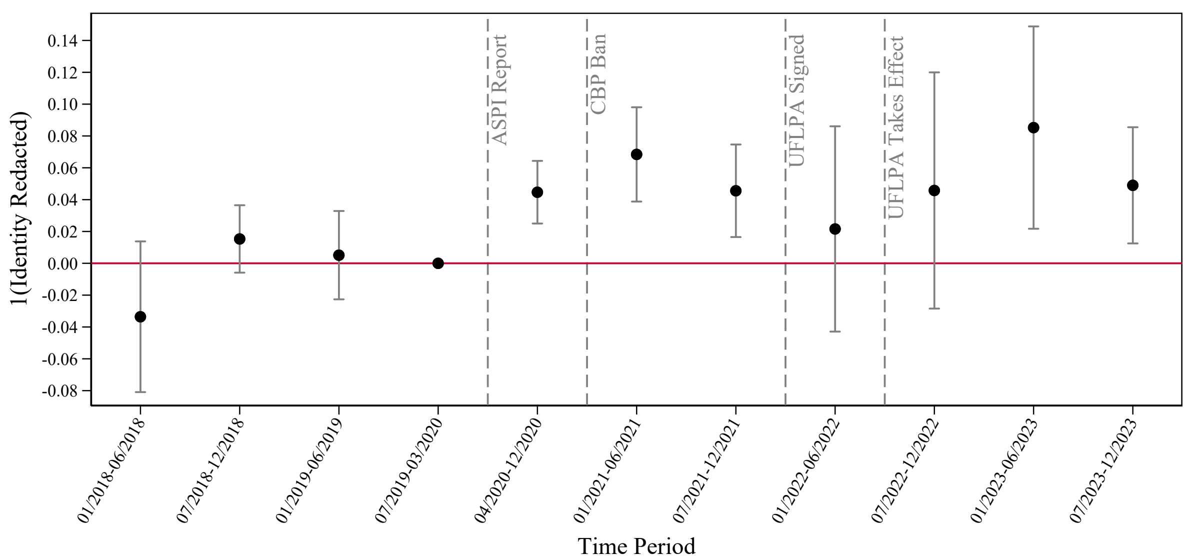


Table 1
Redacted identities by year

This table presents the share of observations (i.e., shipments) that have non-redacted and redacted importer identities by year. Of those observations with redacted importer identities, about 93% also have redacted supplier identities (untabulated).

	Non-redacted		Redacted		Total
Year	Freq.	Row %	Freq.	Row %	Freq.
2013	6,600,593	87.76	920,415	12.24	7,521,008
2014	6,789,236	88.33	896,671	11.67	7,685,907
2015	6,900,693	88.64	884,302	11.36	7,784,995
2016	7,253,871	86.36	1,145,406	13.64	8,399,277
2017	7,588,453	84.29	1,414,665	15.71	9,003,118
2018	8,056,266	83.53	1,587,952	16.47	9,644,218
2019	8,004,620	83.11	1,626,512	16.89	9,631,132
2020	8,199,177	84.22	1,535,731	15.78	9,734,908
2021	9,653,541	81.77	2,151,653	18.23	11,805,194
2022	9,504,673	80.26	2,338,191	19.74	11,842,864
2023	9,057,033	81.78	2,017,891	18.22	11,074,924
Total	87,608,156	84.14	16,519,389	15.86	104,127,545

Table 2

Redacted identities by regions

This table presents the percent of redacted shipments by region (first column) and the overall share of observations in the sample (both redacted and non-redacted) that are associated with that region (second column). In Panel A, we focus on the origin region of the shipment. In Panel B, we focus on the US destination region.

Panel A: Origin region

	% of Regional Shipments Redacted	% of Total Observations from Region
Australia / New Zealand	11.24	0.729
Central Asia	9.863	0.0029
Eastern Asia	17.80	56.59
Eastern Europe	1.971	1.040
Latin America / Caribbean	9.750	7.800
Melanesia	7.889	0.0118
Micronesia	20.74	0.0298
Northern Africa	18.77	0.258
Northern America	11.60	0.262
Northern Europe	9.712	2.357
Polynesia	3.700	0.0229
South-Eastern Asia	17.00	10.27
Southern Asia	14.27	6.185
Southern Europe	13.22	4.919
Sub-Saharan Africa	11.53	0.347
Western Asia	17.54	1.671
Western Europe	13.37	7.509
Total		100

Panel B: Destination region

	% of Regional Shipments Redacted	% of Total Observations to Region
New York Region	17.64	15.09
North Central Region	15.72	11.16
Northeast Region	13.38	5.243
Pacific Region	16.59	35.96
South Central Region	14.23	2.957
Southeast Region	14.54	21.47
Southwest Region	15.26	8.122
Total		100

Table 3
Summary statistics of shipment value and weight

This table provides summary statistics for the value (USD) and weight (metric tons) of imported goods. In Panel A, we present summary statistics for the full sample. In Panels B and C, we restrict the sample to non-redacted (B) and redacted (C) shipments. In Panel D, we present the number of observations in each value-quintile and weight-quintile classification.

Panel A: Full sample

N = 104,127,545	Mean	SD	p10	p25	p50	p75	p90
Value of goods (USD)	167,407	4,167,033	3,900	14,500	37,800	90,000	205,000
Weight, metric tons	55.60	1,445.67	0.51	2.72	9.08	19.15	34.27

Panel B: Non-redacted sample

N = 87,608,156	Mean	SD	p10	p25	p50	p75	p90
Value of goods (USD)	173,247	4,303,869	4,100	14,700	38,600	92,300	210,000
Weight, metric tons	57.29	1,436.62	0.54	2.83	9.23	19.20	34.64

Panel C: Redacted sample

N = 16,519,389	Mean	SD	p10	p25	p50	p75	p90
Value of goods (USD)	136,431	3,348,996	2,900	13,300	33,900	78,300	175,000
Weight, metric tons	46.65	1,492.74	0.39	2.20	8.33	18.79	32.01

Panel D: Observations by value and weight quintiles

Value Quintile	Weight Quintile					
	1	2	3	4	5	Total in Value Quintile
1	16,120,114	3,055,230	484,718	938,839	364,204	20,963,105
Row %	76.90	14.57	2.31	4.48	1.74	
2	3,587,344	7,819,182	4,024,882	2,717,205	2,593,407	20,742,020
Row %	17.30	37.70	19.40	13.10	12.50	
3	770,222	5,384,700	6,565,311	5,425,801	2,644,285	20,790,319
Row %	3.70	25.90	31.58	26.10	12.72	
4	253,504	3,326,057	5,692,510	6,698,404	4,903,501	20,873,976
Row %	1.21	15.93	27.27	32.09	23.49	
5	95,531	1,239,388	4,058,029	5,047,597	10,317,580	20,758,125
Row %	0.46	5.97	19.55	24.32	49.70	
Total in Weight Quintile	20,826,715	20,824,557	20,825,450	20,827,846	20,822,977	*N = 104,127,545*

Table 4

Redacted identities and competitor-based proprietary costs

This table reports results on the relation between competitor-based proprietary costs and redacted identities in shipment records. Panel A provides summary statistics for explanatory variables based on each estimation subsample. We estimate the first regression in Panel B, Column 3 (Equation 1) and the second regression in Panel C, Columns 2 and 4 (Equation 2):

$$\mathbb{1}(IdentityRedacted)_{i,t} = \beta \times \mathbb{1}(ManufacturingGood)_g + \delta_{or,dr,t} + \gamma_{v,w} + \epsilon_{i,t}\ ,$$
$$\mathbb{1}(IdentityRedacted)_{i,t} = \beta \times \mathbb{1}(NewLinkage)_{o,dr,g,t} + \iota_{o,dr,g} + \zeta_{g,t} + \gamma_{v,w} + \epsilon_{i,t}\ .$$

Above, the subscript i reflects a class of shipments with the following characteristics: they originate from country o in region or of the world, have a final destination in region dr of the US, contain goods of type g based on two-digit HS code, and have value in quintile v and weight in quintile w. In Panel B, the sample consists of shipments of manufacturing goods and non-durable goods. In Panel C, the sample consists of shipments after 2016 for supply chain linkages with at least 500 (Columns 1 and 2) or 1,400 (Columns 3 and 4) shipments. Key variables are defined in Appendix A. t-statistics, based on standard errors clustered at the destination region-good type level, are provided in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels (two-tailed), respectively.

Panel A: Summary statistics

	N	Mean	SD	p10	p25	p50	p75	p90
$\mathbb{1}$(Manufacturing Good)	82,089,888	0.59	0.49	0.00	0.00	1.00	1.00	1.00
$\mathbb{1}(\text{New Linkage})_{500}$	78,799,651	0.02	0.14	0.00	0.00	0.00	0.00	0.00
$\mathbb{1}(\text{New Linkage})_{1400}$	76,012,395	0.04	0.19	0.00	0.00	0.00	0.00	0.00

Panel B: Imports of manufacturing goods

$Y = \mathbb{1}$(Identity Redacted)	(1)	(2)	(3)
$\mathbb{1}$(Manufacturing Good)	0.025***	0.026***	0.019**
	(2.85)	(2.95)	(2.24)
Observations	82,089,888	82,089,888	82,089,820
Adj. R^2	0.010	0.012	0.024
Fixed effects:			
Destination region × Year-quarter	Y	Y	
Origin region × Destination region × Year-quarter			Y
Value quintile × Weight quintile		Y	Y

Panel C: Imports of goods indicating new supply chain linkages

$Y = \mathbb{1}$(Identity Redacted)	First 500 shipments		First 1,400 shipments	
	(1)	(2)	(3)	(4)
$\mathbb{1}$(New Linkage)	0.015***	0.011***	0.010***	0.007***
	(3.65)	(4.39)	(2.52)	(2.71)
Observations	78,799,651	78,799,647	76,012,395	76,012,395
Adj. R^2	0.080	0.086	0.075	0.081
Fixed effects:				
Origin country × Destination region × Good type	Y	Y	Y	Y
Year-quarter	Y		Y	
Good type × Year-quarter		Y		Y
Value quintile × Weight quintile		Y		Y

Table 5

Redacted identities and forced labor risks

This table reports results on the relation between origin-country forced labor risks and redacted identities in shipment records. Panel A provides summary statistics for explanatory variables based on each estimation subsample. We estimate the following regression in Panels B and C, Column 2 (Equation 3):

$$\mathbb{1}(IdentityRedacted)_{i,t} = \beta \times \mathbb{1}(ForcedLaborIndicator)_{o} + \phi_{dr,g,t} + \gamma_{v,w} + \epsilon_{i,t} \, .$$

Above, the subscript i reflects a class of shipments with the following characteristics: they originate from country o in region or of the world, have a final destination in region dr of the US, contain goods of type g based on two-digit HS code, and have value in quintile v and weight in quintile w. In Panels B and C, the sample consists of all shipments from origin countries for which the forced labor measures are available. Key variables are defined in Appendix A. t-statistics, based on standard errors clustered at the destination region-good type level, are provided in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels (two-tailed), respectively.

Panel A: Summary statistics

	N	Mean	SD	p10	p25	p50	p75	p90
$\mathbb{1}$(High Vulnerability Country)	103,500,499	0.63	0.48	0.00	0.00	1.00	1.00	1.00
$\mathbb{1}$(Low Government Response Country)	103,733,809	0.77	0.42	0.00	1.00	1.00	1.00	1.00

Panel B: Origin countries with high forced labor vulnerability

$Y = \mathbb{1}$(Identity Redacted)	(1)	(2)
$\mathbb{1}$(High Vulnerability Country)	0.025***	0.019***
	(4.02)	(4.15)
Observations	103,500,499	103,500,369
Adj. R^2	0.009	0.039
Fixed effects:		
Destination region × Year-quarter	Y	
Destination region × Good type × Year-quarter		Y
Value quintile × Weight quintile		Y

Panel C: Origin countries with low government response to forced labor

$Y = \mathbb{1}$(Identity Redacted)	(1)	(2)
$\mathbb{1}$(Low Government Response Country)	0.055***	0.044***
	(7.50)	(8.38)
Observations	103,733,809	103,733,678
Adj. R^2	0.012	0.040
Fixed effects:		
Destination region × Year-quarter	Y	
Destination region × Good type × Year-quarter		Y
Value quintile × Weight quintile		Y

Table 6

Redacted identities and increases in forced labor scrutiny

This table reports results on the relation between events related to increased forced labor scrutiny (see Section 2.3) and redacted identities in shipment records. In Panel A, we present summary statistics for our estimation sample of cotton/apparel shipments and manufacturing/durable shipments from 2018 to 2023. We estimate the following difference-in-difference-in-differences design in Panel B, Column 2 (Equation 4):

$$\mathbb{1}(IdentityRedacted)_{i,t} = \beta \times [\mathbb{1}(Cotton, Apparel)_a * \mathbb{1}(FromChina)_c * PostASPI_t] + \alpha_{a,c} + \kappa_{a,t} + \varphi_{c,t} + \epsilon_{i,t} .$$

Above, the subscript i reflects a class of shipments with the following characteristics: they originate from country o in region or of the world, have a final destination in region dr of the US, contain goods of type g based on two-digit HS code, and have value in quintile v and weight in quintile w. a (Cotton, Apparel (Y/N)) indicates whether a good type is (or is not) cotton/apparel, and c (From China (Y/N)) indicates whether the shipment of that good originates (or does not originate) from China. Key variables are defined in Appendix A. t-statistics, based on standard errors clustered at the destination region-good type level, are provided in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels (two-tailed), respectively.

Panel A: Summary statistics

N = 45,645,685	Mean	SD	p10	p25	p50	p75	p90
𝟙(Identity Redacted)	0.19	0.39	0.00	0.00	0.00	0.00	1.00
𝟙(Cotton, Apparel)	0.12	0.32	0.00	0.00	0.00	0.00	1.00
𝟙(From China)	0.41	0.49	0.00	0.00	0.00	1.00	1.00
PostASPI	0.67	0.47	0.00	0.00	1.00	1.00	1.00
𝟙(Cotton, Apparel) × 𝟙(From China)	0.04	0.20	0.00	0.00	0.00	0.00	0.00
𝟙(Cotton, Apparel) × 𝟙(From China) × PostASPI	0.03	0.16	0.00	0.00	0.00	0.00	0.00

Panel B: Redacted identities after intensified forced labor scrutiny

Y = 𝟙(Identity Redacted)	(1)	(2)	(3)	(4)
𝟙(Cotton, Apparel) × 𝟙(From China)	0.016			
	(1.27)			
PostASPI	0.012***			
	(4.17)			
𝟙(Cotton, Apparel) × 𝟙(From China) × PostASPI	0.034*	0.052***	0.048***	0.049***
	(1.90)	(3.02)	(2.67)	(2.73)
Observations	45,645,685	45,645,685	45,645,685	45,645,685
Adj. R^2	0.001	0.008	0.026	0.027
Fixed effects:				
Cotton, Apparel (Y/N) × From China (Y/N)		Y		
Cotton, Apparel (Y/N) × Year-quarter		Y	Y	Y
From China (Y/N) × Year-quarter		Y	Y	Y
Destination region × Year-quarter			Y	Y
Good type × From China (Y/N) × Destination region			Y	Y
Value quintile × Weight quintile				Y

Internet Appendix

Redacted Identities in Shipment Records: Evidence from Forced Labor Scrutiny in Supply Chains

Sandra G. Schafhäutle and Gurpal S. Sran

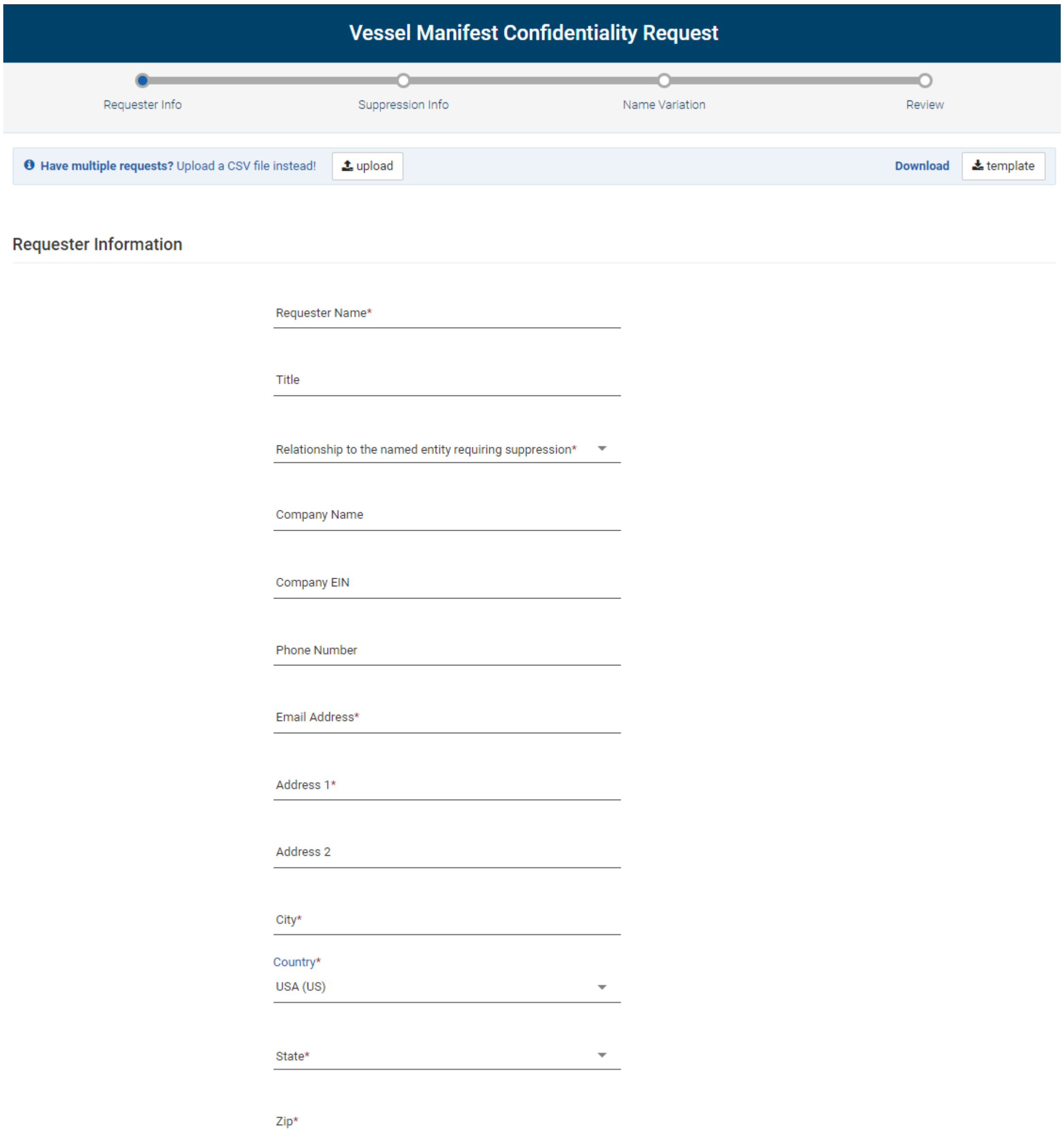
February 2025

Figure IA.1

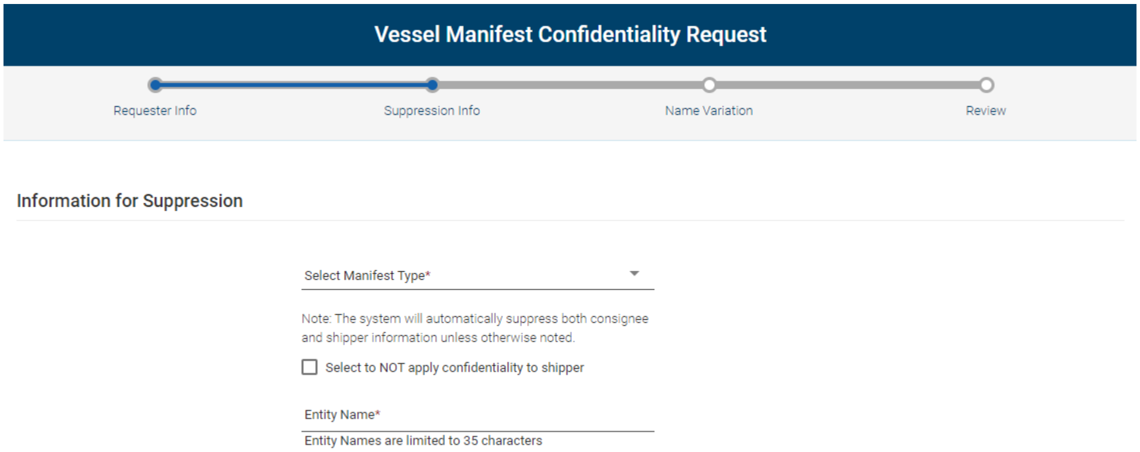
Requesting manifest confidentiality

This figure provides screenshots of the current online application to request manifest confidentiality (see www.cbp.gov). There is a three-step process of filling out the online application.

Step 1: Importers must fill out identifying information



Step 2: Importers input their entity name to be redacted and agree to automatically redact the identifying information of all suppliers unless they opt-out



Step 3: Importers can list any (or upload a CSV) of their name variants that might appear on incoming shipments such that redaction is more complete

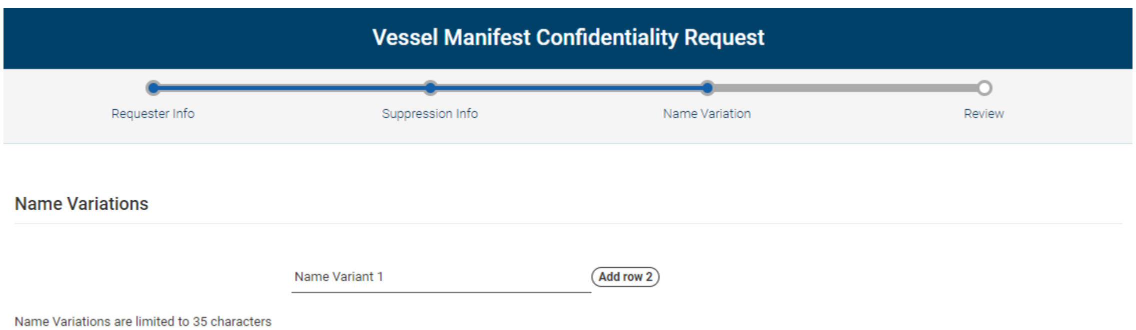


Table IA.1

Robustness tests

This table reports robustness tests related to our results in Table 6. Unlading region is the US geographical area of the port where a shipment is unloaded from the vessel. t-statistics, based on standard errors clustered at the destination region-good type level (unless indicated otherwise), are provided in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels (two-tailed), respectively.

	β	Observations	Adj. R^2
Baseline specification: Table 6, Panel B, Column 2	0.052***	45,645,685	0.008
	(3.02)		
Include Cotton, Apparel (Y/N) × Origin country FEs	0.047***	45,645,669	0.033
	(2.80)		
Include Unlading region × Year-quarter FEs	0.050***	45,645,685	0.009
	(2.90)		
Extend pre-period to 2017	0.045***	52,095,981	0.008
	(3.08)		
Unrestricted control group	0.044**	63,733,240	0.010
	(2.51)		
Exclude shipments originating from India	0.049***	43,471,360	0.007
	(2.66)		
Define geographical treatment as $\mathbb{1}$(From Eastern or South-Eastern Asia)	0.068***	45,645,685	0.008
	(3.67)		
Collapse to From China (Y/N)-good type-year quarter, add interactive good type FEs and cluster by good type	0.076**	2,064	0.711
	(2.11)		
Cluster standard errors by good type	0.052**	45,645,685	0.008
	(2.53)		
Cluster standard errors by origin country	0.052***	45,645,685	0.008
	(2.74)		
Cluster standard errors by origin country-good type	0.052***	45,645,685	0.008
	(2.69)		

Table IA.2

Disappearance of importer trade names and increases in forced labor scrutiny

This table reports results on the relation between events related to increased forced labor scrutiny (see Section 2.3) and the disappearance of specific importer trade names from Panjiva data (i.e., an alternate proxy for identity redactions). In this analysis, we collapse the transaction-level data to the trade name level and then restrict the sample to trade names associated with a large number of shipments in the pre-ASPI period (i.e., those trade names are equal to or above the 90th percentile of the distribution of the number of shipments from 2018Q1 through 2020Q1). Using this baseline sample, we estimate the following designs in Panels A and B, respectively:

$$\mathbb{1}(\textit{Disappeared})_j = \beta \times \mathbb{1}(\textit{Exposed})_j + \alpha_s + \epsilon_j \ ,$$
$$\mathbb{1}(\textit{Disappeared})_j = \beta \times \textit{Exposure}_j + \alpha_s + \epsilon_j \ .$$

Above, the subscript j reflects a trade name (as identified by Panjiva's unique consignee identifier). s is a categorical variable (30 quantiles) representing the number of pre-ASPI shipments for each j, and α_s represents size-quantile fixed effects (s and α_s are calculated for each estimation sample separately). $\mathbb{1}(\textit{Exposed})_j$ is an indicator variable equal to one if 50% or more of the trade name's pre-ASPI shipments were cotton and apparel shipments originating from China, and zero otherwise. $\textit{Exposure}_j$ is the underlying continuous variable. $\mathbb{1}(\textit{Disappeared})_j$ is an indicator variable equal to one if trade name j has *no* shipment records *post*-ASPI, and zero otherwise. In Column 1, we report estimates for the baseline sample, and in Column 2, we add size-quantile fixed effects. In Column 3, we condition the sample to trade names that primarily source goods from China during the pre-ASPI period. In Column 4, we further condition the sample to trade names associated with at least one cotton and apparel shipment during the pre-ASPI period. t-statistics, based on standard errors clustered at the trade name level, are provided in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels (two-tailed), respectively.

Panel A: Disappearance of trade names, binary exposure variable

$Y = \mathbb{1}$(Disappeared)	(1)	(2)	(3)	(4)
$\mathbb{1}$(Exposed)	0.085***	0.081***	0.073***	0.063***
	(8.71)	(8.34)	(7.45)	(5.96)
Constant	0.053***			
	(48.04)			
Observations	42,366	42,366	14,829	5,109
Adj. R^2	0.004	0.012	0.013	0.022
Mean, $\mathbb{1}(\textit{Exposed})$	0.030	0.030	0.085	0.248
Fixed effects:				
Size quantile		Y	Y	Y

Panel B: Disappearance of trade names, continuous exposure variable

$Y = \mathbb{1}$(Disappeared)	(1)	(2)	(3)	(4)
Exposure	0.110***	0.105***	0.093***	0.084***
	(9.60)	(9.28)	(7.70)	(6.16)
Constant	0.052***			
	(46.38)			
Observations	42,366	42,366	14,829	5,109
Adj. R^2	0.005	0.013	0.014	0.024
Mean, *Exposure*	0.035	0.035	0.085	0.247
Fixed effects:				
Size quantile		Y	Y	Y