

Selection into Tariff Avoidance and the Mismeasurement of Tariff Incidence

Jackson Mejia*

A large literature suggests that the incidence of tariffs falls almost entirely on the United States. I argue this finding is partly explained by a selection bias. When costly tariff avoidance is possible, low-passthrough exporters—those that would otherwise bear the most incidence—are the most likely to undertake avoidance strategies that remove them from common datasets. Their exit biases common regression specifications estimates upward, overstating the incidence borne by the United States. I quantify this effect using transshipment as a key, measurable avoidance channel. Applying a methodology from Freund (2024) to Comtrade data, I estimate that accounting for this selection reduces measured U.S. incidence from near 100% to a range of 60% to 90%. The theory also reveals a novel enforcement margin. When imports are fully tax-deductible, as with intermediate and capital goods, the avoidance margin is less profitable and transshipment is lower.

*Massachusetts Institute of Technology, jpmejia@mit.edu. First Draft: September 2025. I am especially grateful to Steve Miran and Kim Ruhl for extensive discussions and comments on the development of these ideas, as well as to Maxx Cook, Andrew Johnston, and Pedro Martinez-Bruera for useful conversations and comments. The views and conclusions contained in this document are mine and should not be interpreted as representing the official policies, either expressed or implied, of any government agency.

1 Introduction

Following the onset of the 2018 trade war between China and the United States, a battery of papers documented that the latter’s importers bore the full incidence of tariffs (Amiti, Redding, and Weinstein 2019; Cavallo et al. 2021; Fajgelbaum et al. 2020; Flaaen, Hortaçsu, and Tintelnot 2020). This finding is puzzling. Given the large size of the U.S. economy, theory predicts that incidence should have been shared with China. This paper argues that this puzzle can be partly explained by a selection bias driven by a critical, but often overlooked, behavioral response: costly tariff avoidance. Specifically, exporters can reroute goods through a third country to evade tariff—a practice known as transshipment. Empirically, this practice accounts for a large share of imports nominally from countries like Mexico and Vietnam that actually originate from China (Freund 2024; Iyoha et al. 2025). This rerouting creates two problems: the government loses tariff revenue, and our measurement of incidence on the remaining trade becomes biased upward.

I propose a framework in which monopolistically competitive firms with heterogeneous marginal costs choose whether to ship directly and bear some of the tariff incidence or pay a fixed cost to transship. The key insight is that this decision is not random. Exporters that are hurt most by the tariff have the strongest incentive to pay the fixed cost to avoid it. When markups are heterogeneous, like with Kimball (1995) demand, the firms which select into transshipment have the highest markups and the lowest marginal costs.¹ Because these low-passthrough firms selectively exit the direct-shipping sample, the firms that remain are disproportionately those that pass are able to pass the tariff’s cost onto U.S. importers.

The selection creates a bias in empirical estimates of tariff passthrough. “Border designs,” which rely on customs unit values or aggregate import price indices constructed from direct shipments, are susceptible because transshippers are unobserved and drop out of the sample. This approach is used in influential papers like Amiti, Redding, and Weinstein (2019), Fajgelbaum et al. (2020), and parts of Cavallo et al. (2021). In contrast, “cohort designs” that track the same product or firm from origin to destination are not subject to selection bias. This includes the retailer tracking in parts of Cavallo et al. (2021) and the analysis in Flaaen, Hortaçsu, and Tintelnot (2020).

1. We know from a large number of studies (Gopinath and Itskhoki 2010; Amiti, Itskhoki, and Konings 2019; De Loecker, Eeckhout, and Unger 2020; Edmond, Midrigan, and Xu 2023) that markups are heterogeneous, so this is a weak condition. Whether heterogeneous foreign markups are good or bad for domestic welfare depends on the welfare components under consideration (consumer surplus, tariff revenue, and terms-of-trade). The selection mechanism here reduces measured border pass-through and increases compliant flows; the net welfare effect is therefore ambiguous ex ante and depends on how these components trade off in the application at hand.

The bias in border designs is empirically large. Using a methodology from Freund (2024) to identify transshipment in U.N. Comtrade data, I estimate that between 10-40% of goods targeted by the 2018 tariffs on China were transshipped. Theory provides a direct way to translate this avoidance share into a correction for measurement bias. My estimates imply that the true tariff passthrough is between 60-90%, substantially below the 100% documented by the border-design literature. Since my own estimates do not account for other forms of avoidance, they are also likely too large.²

Such large-scale avoidance is not merely a measurement problem. It also creates a significant welfare problem, which can be analyzed through the lens of the tariff's marginal cost of public funds (MCPF) (Finkelstein and Hendren 2020), which is roughly

$$\text{Tariff MCPF} \approx \frac{\text{Domestic Incidence}}{\text{Fiscal Externality}}$$

The MCPF is the welfare cost of raising an additional dollar of government revenue and combines two elements: the domestic incidence of the tax, and the fiscal externality from the behavioral response. This paper makes a contribution to understanding both elements. My finding that true incidence is lower than previously thought directly corrects measurement of the numerator, implying that the welfare cost of the tariff is smaller than the literature suggests. At the same time, the large volume of transshipment constitutes the denominator's fiscal externality. This raises a crucial enforcement question: what policy tools can shrink the behavioral response and thereby lower the overall welfare cost of tariffs?

This paper identifies a powerful but overlooked enforcement margin embedded within the domestic tax system itself. While the public finance literature focuses on enforcement tools like audits, inspections, and penalties (Allingham and Sandmo 1972; Slemrod and Yitzhaki 2002; Bhandari et al. 2024; Slemrod 2019), I show theoretically that allowing importer firms to fully deduct the cost of imported intermediate and investment goods is isomorphic to such measures. Deductibility lowers the importer's effective burden of the tariff, making them less sensitive. This allows exporters to pass through more of the tariff on compliant shipments. That, in turn, raises the profitability of direct shipping relative to transshipment and thus reducing avoidance. This insight has direct policy relevance because the One Big Beautiful Bill Act made expensing permanent contemporaneously with the wave of new tariffs on many foreign countries.

2. This rerouting is a salient example of creative (but costly) avoidance strategies to which our selection mechanism applies, including misreporting value to fall under de minimis thresholds (Fajgelbaum and Khandelwal 2024), breaking up shipments into smaller parcels, or misclassifying them (Fisman and Wei 2004).

2 Baseline Model

This section develops a partial equilibrium model that serves two purposes. First, it formalizes the selection mechanism through which costly avoidance leads to an upward bias in measured tariff passthrough. Second, it provides the framework to analyze how other components of tax policy interact with routing choice to act as an enforcement margin.

2.1 Environment.

Consider a static partial equilibrium setting with monopolistic competition over differentiated varieties indexed by $i \in \mathcal{J}$. Each exporter sets a price p_i and sells to a Home market that imposes a uniform ad valorem tariff $\tau \geq 0$ on direct shipments from the targeted origin. The delivered price paid in home is $\tilde{p}_i^D = (1 + \tau)p_i$. The firm may pay a fixed cost $F > 0$ to avoid the tariff. In practice, this would resemble a Chinese exporter shipping through Vietnam to export a good to the United States. In that case $\tilde{p}_i^T = p_i$. We denote the choices as $c \in \{D, T\}$, where D denotes direct shipping and T denotes transshipping.

The fixed cost F captures the various expenses and risks associated with establishing and using a transshipment route. In reality, this cost is not a single number but is likely determined by several factors. These include the logistical costs tied to the distance and infrastructure of the intermediary country, the characteristics of the product itself (e.g., electronics are easier to reroute than fresh produce or heavy machinery), and the legal or administrative fees required to relabel goods. While F likely varies across firms and products, this model assumes a uniform cost for tractability. Acknowledging this heterogeneity, however, suggests that the incentive to transship is strongest for firms that produce goods that are easy to reroute and have access to efficient transshipment hubs.

Generalized demand. We impose minimal assumptions on Home demand to nest several important cases. Demand for each variety depends on its own delivered price and an aggregator summarizing the prices of all other varieties. Formally, the demand for variety i takes the form

$$q_i = D_i(\tilde{p}_i \mid \mathbf{P}), \quad \mathbf{P} \equiv \mathbf{P}(\mathbf{P}_{-i}; \Theta),$$

where \mathbf{P} is a pricing aggregator parameterized by Θ , and a sufficient statistic for the competitive environment reflected by firm i . Importantly, it nests CES and Kimball demand, among others. \mathbf{P} is exogenous at the firm level. This corresponds to a standard monopolistic competition setup in which each firm is small relative to the market and takes the

price distribution of rivals as given. Importantly, we index the elasticity of demand for each variety by ε and its superelasticity by κ . Under CES, $\kappa = 0$, while Kimball has $\kappa < 0$. That is, elasticity increases with price. Given that, we impose several regularity conditions on demand:

Assumption 1 (Demand). *The demand for each variety i depends on its own delivered price \tilde{p}_i and on a price aggregator \mathbf{P} summarizing rivals. Fixing \mathbf{P} as parametric at the variety level, the primitives satisfy:*

D1. Downward sloping own demand and outward shifts in rivals: $\frac{\partial D_i}{\partial \tilde{p}_i} < 0$ and $\frac{\partial D_i}{\partial \mathbf{P}} > 0$.

D2. Smoothness: $D_i(\cdot \mid \mathbf{P}) \in C^2$ in own price.

D3. Elastic demand: $\varepsilon_i(\tilde{p}_i \mid \mathbf{P}) \equiv -\frac{\tilde{p}_i}{D_i} \frac{\partial D_i}{\partial \tilde{p}_i} > 1$.

D4. Profit concavity / local stability: For either channel $c \in \{D, T\}$ with delivered price \tilde{p}_i^c ,

$$\varepsilon(\tilde{p}_i^c \mid \mathbf{P}) - 1 - \kappa(\tilde{p}_i^c \mid \mathbf{P}) > 0.$$

D5. Kimball curvature: $\kappa_i(\tilde{p}_i \mid \mathbf{P}) \equiv -\frac{\partial \ln \varepsilon_i(\tilde{p}_i \mid \mathbf{P})}{\partial \ln \tilde{p}_i} \leq 0$ (equals 0 under CES). We also impose two extra conditions:

$$(a) \frac{\partial \kappa(\tilde{p}_i \mid \mathbf{P})}{\partial \ln \tilde{p}_i} \leq 0.$$

$$(b) \kappa'(\tilde{p}) \equiv \partial \kappa(\tilde{p}) / \partial \ln \tilde{p} \geq -\frac{\varepsilon(\tilde{p})}{\varepsilon(\tilde{p})-1} \kappa(\tilde{p})^2.$$

D1 ensures that demand slopes down in own delivered price and that higher rival prices (the aggregator \mathbf{P}) shift out demand. D2 gives smoothness for envelope/implicit-function arguments, and D3 ($\varepsilon > 1$) guarantees an interior Lerner markup. D4 imposes the second-order condition in elasticity form, $\varepsilon - 1 - \kappa > 0$, equivalent to strict profit concavity at the optimum. D5 ($\kappa \leq 0$) nests CES as $\kappa \equiv 0$ and the empirically relevant Kimball class where elasticity rises with price. For some results we impose the mild strengthening D5a–b. Together they imply that survivor pass-through is (weakly) increasing in delivered price and, because p^* increases in m (see Lemma 1), also (weakly) increasing in marginal cost.

Firm problem. Each firm $i \in \mathcal{I}$ has a constant marginal cost m_i . Exporter marginal costs m_i are independently drawn from a continuous distribution with support $[\underline{m}, \bar{m}] \subset (0, \infty)$ and density $f_m(m) > 0$. The distribution is independent of policy τ and of the aggregator \mathbf{P} , and has no atoms, so any cutoff \hat{m} is unique and yields smooth comparative statics.

Each firm must choose whether to pay a uniform fixed cost F to avoid the tariff via transshipment or absorb some part of the tariff and ship directly to the home nation. Given $c \in \{D, T\}$, the firm solves

$$\pi_i(m_i; c) = \max_{p_i \geq m_i} \underbrace{(p_i - m_i) D_i(\tilde{p}_i^c | \mathbf{P})}_{\equiv \Phi_i(p_i; m_i, \tilde{p}_i^c)} - \mathbb{1}\{c = T\}F, \quad \tilde{p}_i^D = (1 + \tau)p_i, \tilde{p}_i^T = p_i.$$

Lemma 1 (Lerner condition and monotone comparative statics). *Under D1–D4, any $p_i^*(m_i; c)$ is interior and satisfies*

$$\frac{p_i^* - m_i}{p_i^*} = \frac{1}{\varepsilon_i(\tilde{p}_i^c | \mathbf{P})'}, \quad \tilde{p}_i^c \in \{(1 + \tau)p_i^*, p_i^*\}.$$

Moreover: (i) p_i^* strictly increases in m_i ; (ii) in the direct channel D , p_i^* is weakly decreasing in τ (constant under CES); and (iii) q_i^* strictly decreases in m_i and, for D , in τ .

Proof: Appendix C.1

Lemma 1 simply establishes the standard Lerner condition, namely that firms will set their markup inversely to the negative elasticity of demand facing the firm. This makes all comparative statics run through how the tariff shifts demand and how elasticity varies with price. Part (i) says that higher marginal cost pushes the firm's optimal price upward, giving us a clean monotone ordering in m that will pin down the cutoff. Part (ii) highlights that the tariff is a demand shifter: under CES, elasticity is constant, so the firm's optimal markup does not change. Under Kimball-type demand, elasticity itself rises with delivered prices, so firms temper the delivered-price increase by lowering (or raising less) their price; this is the source of heterogeneous passthrough across i . Part (iii) records the quantity implications: higher m or higher τ cuts q_i^* , which is exactly why direct-shipment profits fall relative to avoidance as τ rises. These three properties together drive selection into transshipment and guarantee a unique, upward-moving cutoff in m , which I formalize subsequently.

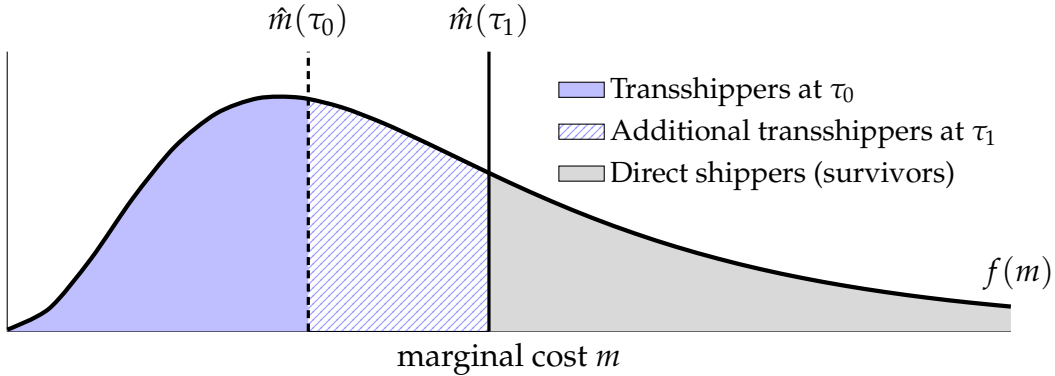
2.2 The Routing Choice

Before formalizing the cutoff, it is helpful to lay out the core intuition for why certain firms select into avoidance. The model features two key elements: firms have heterogeneous marginal costs, and under Kimball demand, tariff passthrough rises with marginal cost. This means that low-cost firms, which operate with higher markups, have low passthrough. A low passthrough rate means these high-markup firms must absorb a

larger share of the tariff themselves, cutting directly into their profit margins. This “price effect” hurts their operating profits more severely than the simple “quantity effect” (lost sales volume) faced by firms that can fully pass on the tariff. Therefore, it is precisely these low-cost, high-markup firms, the ones who would otherwise have to absorb the tariff, that have the strongest incentive to pay the fixed cost to avoid it altogether. Higher-cost firms, which can pass on most of the tariff, find it more profitable to continue shipping directly. This selection process is the key mechanism driving the paper’s results.

Given that intuition, we now formally characterize how firms sort between direct shipment and transshipment. The only asymmetry across channels is that the tariff enters the delivered price in the direct channel—buyers face $(1 + \tau)p$ —while the avoidance benchmark prices against $D(p)$ but incurs a fixed cost F . Given our primitives **D1–D4**, the direct and avoidance pricing problems are well behaved, so we can study routing by comparing their optimized values as functions of marginal cost m and the tariff.

Figure 1: An increase in tariffs raises the cutoff



Note: Distribution of marginal costs and tariff-induced selection. The cutoff $\hat{m}(\tau)$ partitions the cohort: firms with $m < \hat{m}(\tau)$ avoid (blue), while $m \geq \hat{m}(\tau)$ ship direct (gray). A higher tariff shifts the cutoff right, increasing the transshipment share θ and shrinking the survivor set.

Formally, define the channel values

$$V(m, 1 + \tau) \equiv \max_{p \geq m} (p - m) D((1 + \tau)p)$$

$$\pi^D(m) = V(m, 1 + \tau), \quad \pi^T(m) = V(m, 1) - F,$$

and the relative gain from avoidance

$$\Delta(m; \tau) \equiv \pi^T(m) - \pi^D(m) = V(m, 1) - V(m, 1 + \tau) - F.$$

$\Delta(m; \tau)$ is the gain from transshipping for a firm with marginal cost m . It falls in m and

risers in τ .

Proposition 1 (Cutoff routing and comparative statics). *Under D1–D4:*

(a) (Monotone selection in cost) $\Delta(m; \tau)$ is strictly decreasing in m . Hence there exists a unique cutoff $\hat{m}(\tau) \in [\underline{m}, \bar{m}]$ such that

$$m < \hat{m}(\tau) \Rightarrow \text{Avoid}, \quad m \geq \hat{m}(\tau) \Rightarrow \text{Direct}.$$

(b) (Tariff comparative static) $d\hat{m}/d\tau > 0$. Consequently $S_i \equiv \mathbb{1}\{m_i \geq \hat{m}(\tau)\}$ is (weakly) decreasing in τ , and the transshipment share $\theta(\tau) \equiv \Pr(S_i = 0)$ is (weakly) increasing in τ .

Proof. Appendix C.2.

The threshold $\hat{m}(\tau)$ partitions the cohort into two routing groups: firms with $m < \hat{m}(\tau)$ optimally pay F and avoid, while firms with $m \geq \hat{m}(\tau)$ continue to ship direct. The reason is straightforward. As m rises, the firm's per-unit margin $(p^* - m)$ shrinks, equilibrium quantity falls, and overall profits decline in both channels. But the decline is not symmetric. The direct channel's value $V(m, 1 + \tau)$ is hit harder because the tariff raises the delivered price $(1 + \tau)p$ and suppresses demand. The avoidance benchmark $V(m, 1)$ faces no such wedge. Thus for low-cost, high-volume firms, the tariff cuts deeply into total operating profits, making it worthwhile to pay the fixed cost and avoid. For higher-cost firms, volumes are small and the incremental tariff burden is limited; these firms can cover themselves by passing much of the tariff into delivered prices. For them the one-time cost F is not worth paying, so they remain in the direct channel. This monotone ordering in m generates a unique cutoff and explains why $\hat{m}(\tau)$ shifts right when the tariff rises.

Since a higher tariff pushes low-markup firms into transshipment, the remaining firms are, on average, better able to pass through the tariff onto domestic importers. I now turn to examining the corresponding pricing response.

2.3 Passthrough

We now characterize the pricing response to a tariff for firms that continue to ship direct. For a survivor i , holding \mathbf{P} fixed,

$$\beta_i \equiv \frac{d \ln \tilde{p}_i^D}{d\tau}, \quad \tilde{p}_i^D = (1 + \tau) p_i^*(m_i; D).$$

Writing the Lerner condition at the delivered price and differentiating with respect to $\ln(1 + \tau)$ yields the expression

$$\beta_i = \frac{1}{1 + \tau} \cdot \frac{\varepsilon((1 + \tau)p_i^*) - 1}{\varepsilon((1 + \tau)p_i^*) - 1 - \kappa((1 + \tau)p_i^*)}, \quad (1)$$

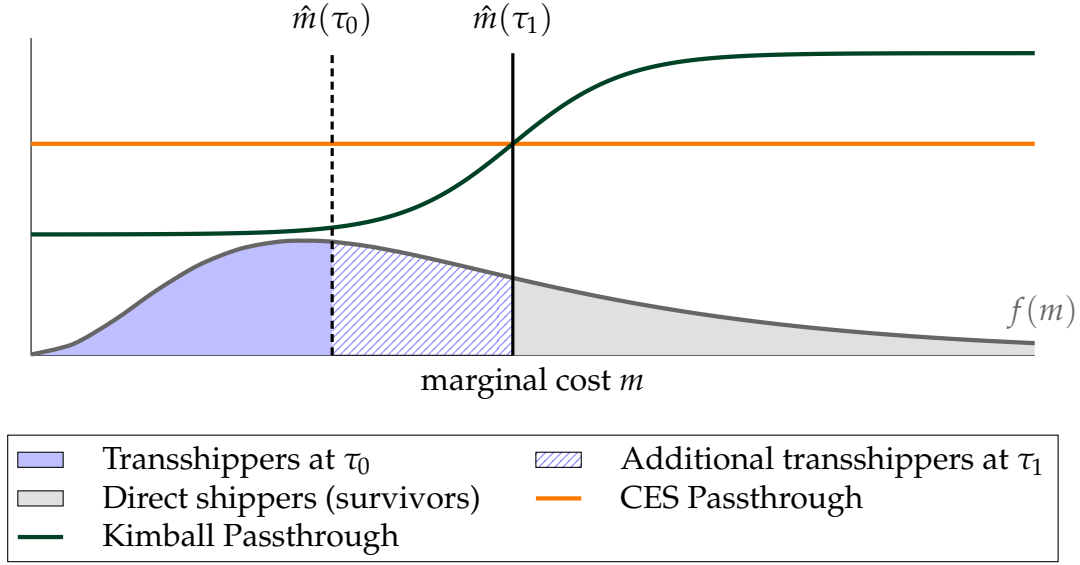
where $\varepsilon(\cdot)$ is the own-price elasticity and $\kappa(\cdot) \equiv -\partial \ln \varepsilon(\cdot) / \partial \ln(\cdot)$ is the superelasticity, both evaluated at the delivered price $(1 + \tau)p_i^*$. The denominator is positive under **D4**. In the CES benchmark ($\kappa \equiv 0$), passthrough is homogeneous: $\beta_i = (1 + \tau)^{-1}$ for all survivors.

Proposition 2 (CES vs. Kimball among survivors). *Fix P and assume D1–D4. (i) CES: If demand is CES ($\kappa \equiv 0$), then $\partial p_i^* / \partial \tau = 0$ and $\beta_i = (1 + \tau)^{-1}$ for every survivor i . (ii) Kimball demand: If, in addition, D5a–b holds, then $\partial p_i^* / \partial \tau \leq 0$ and β_i is (weakly) increasing in m_i within the survivor set.*

Equation (1) shows that passthrough is the product of a mechanical factor $1/(1 + \tau)$ and a curvature factor $(\varepsilon - 1)/(\varepsilon - 1 - \kappa)$. Under CES, elasticity is constant, the optimal markup does not change with τ , and all survivors load the same fraction of the tariff into the delivered price. With Kimball demand, that is not the case. Heterogeneous markups imply that passthrough rises with marginal cost. Low marginal cost firms are less able to pass costs on to domestic consumers, so they choose to transship. From Proposition 1, that also means when tariffs rise, the cutoff shifts right and low- m firms exit. Under Kimball demand the Lerner markup falls with delivered price, so the average markup of survivors falls, even as their average passthrough rises.

Figure 2 illustrates the mechanism. We take the same distribution as Figure 1 and overlay two lines: a flat tangerine line corresponding to homogeneous markups under CES demand, and an upward sloping green line corresponding to heterogeneous markups under Kimball demand. Monotonicity in the superelasticity guarantees that passthrough is weakly increasing among survivors. At any given tariff, survivors are the high- m types with (weakly) higher passthrough; as τ rises, selection tilts the survivor set toward these high- β types even though each firm's β can fall mechanically with τ . That implies something important about measured passthrough.

Figure 2: Passthrough $\beta(m)$ under CES and Kimball



Note: The density $f(m)$ is split by the routing cutoffs $\hat{m}(\tau_0)$ (dashed) and $\hat{m}(\tau_1)$ (solid): transshippers are blue, additional transshippers hatched, survivors gray. Overlaid are $\beta(m)$ under CES (flat, tangerine) and Kimball (increasing, green), scaled to share the same mean $\bar{\beta}$. With Kimball, the survivor mean $\bar{\beta}$ shifts up when τ rises because the cutoff prunes low- β types.

3 Selection in Passthrough Regressions

Theory indicates that heterogeneous markups implies selection into who transships. This has important empirical implications for how we measure passthrough because many regressions of prices on tariffs only observe the prices of direct survivors, which yields some degree of bias in passthrough regressions. In this section, I formally discuss that and provide an estimate for the degree of bias.

3.1 The Bias in Border Designs

Border prices are only observed for firms that continue to ship direct after the tariff change. Let $S_i \in \{0, 1\}$ be the survivor indicator in the post-period ($S_i = 1$ if i ships direct, 0 if i avoids), and let β_i denote firm i 's direct-channel pass-through. Consider a regression run on data on imported goods from the targeted country

$$\Delta \ln \tilde{p}_i = \alpha + \beta \Delta \tau + \varepsilon_i, \quad (2)$$

where $\tilde{p}_i = (1 + \tau)p_i$ is the delivered price and $\Delta\tau$ is the policy change.³ Since $\Delta \ln \tilde{p}_i = \beta_i \Delta\tau$ for a direct shipper i , the OLS coefficient in (2) run on survivors identifies the survivor average:

$$\hat{\beta}_{\text{border}} = \mathbb{E}[\beta_i \mid S_i = 1]$$

Our cohort object of interest is $\beta_{\text{cohort}} \equiv \mathbb{E}[\beta_i]$, the average direct-channel pass-through for the pre-tariff cohort. The following identity makes the selection bias transparent.

Proposition 3 (Border vs cohort bias). *Let $\Pr(S_i = 1)$ be the survivor share. Then*

$$\beta_{\text{border}} \equiv \mathbb{E}[\beta_i \mid S_i = 1] = \underbrace{\mathbb{E}[\beta_i]}_{\beta_{\text{cohort}}} + \frac{\text{Cov}(\beta_i, S_i)}{\Pr(S_i = 1)}. \quad (3)$$

Proof. $\mathbb{E}[\beta_i \mid S_i = 1] = \mathbb{E}[\beta_i S_i] / \mathbb{E}[S_i] = (\mathbb{E}[\beta_i] \mathbb{E}[S_i] + \text{Cov}(\beta_i, S_i)) / \Pr(S_i = 1)$. \square

Under D5 and Proposition 1, β_i and S_i are both increasing in m , so $\mathbb{E}[\beta_i \mid S_i = 1] \geq \mathbb{E}[\beta_i \mid S_i = 0]$ and $\beta_{\text{border}} \geq \beta_{\text{cohort}}$. Two immediate corollaries follow. First, under CES ($\kappa \equiv 0$), all survivors have the same firm-level passthrough $\beta_i = (1 + \tau)^{-1}$, so $\text{Cov}(\beta_i, S_i) = 0$ and $\beta_{\text{border}} = \beta_{\text{cohort}}$, resulting in no selection bias. By contrast, under Kimball-type demand with D5 and the single-crossing routing result (Proposition 1), β_i is increasing in m_i and the survivor indicator S_i is also increasing in m_i ; hence $\text{Cov}(\beta_i, S_i) > 0$ and $\beta_{\text{border}} > \beta_{\text{cohort}}$. Intuitively, the border regression averages passthrough over the *selected* survivor set; curvature makes β_i rise with cost, the cutoff prunes low- m (low- β) firms, and the survivor mean shifts up. On the other hand, with CES homogeneity, selection does not tilt the average. This is simply an empirical description of what we visualized in Figure 2. Since markups are heterogeneous in practice (Edmond, Midrigan, and Xu 2023), that means standard passthrough regressions likely overstate the degree to which incidence is borne by domestic consumers.⁴

When is there selection bias? The selection term in (3) applies whenever the outcome used in the regression is observed only for *direct shippers* or is an aggregate constructed

3. In practice, most empirical work relies on much richer specifications, with variation across units in a number of dimensions. That is irrelevant for the broader point about selection.

4. Note that this has implications for aggregate measurement as well. For domestic incidence, the relevant object is the cohort passthrough, which maps the tariff into the price index for the exposed import bundle faced by residents. For a small change $d\tau$ on an ad valorem tariff and expenditure share s_M of the exposed set,

$$d \ln \text{CPI}_M \approx s_M \beta_{\text{cohort}} d\tau,$$

so any CPI or real-income calculation that is linear in passthrough should be based on β_{cohort} , not the border estimate.

from their transactions. In those designs the estimating sample is $S_i = 1$ by construction, so the OLS estimand is $\beta_{\text{border}} = \mathbb{E}[\beta_i \mid S_i = 1]$ and the gap to the cohort object $\beta_{\text{cohort}} = \mathbb{E}[\beta_i]$ equals $\text{Cov}(\beta_i, S_i) / \Pr(S_i = 1)$.

When firm–product transactions are followed across routing so the *cohort is fixed* at the micro level, there is no selection bias in estimating the cohort passthrough. This requires observing the same exporter–importer (or firm–product–destination) unit before and after the tariff and keeping it in the sample whether it ships direct or avoids. Parts of Cavallo et al. (2021) implement this item-level tracking. Similarly, Flaaen, Hortaçsu, and Tintelnot (2020) avoid this bias by focusing solely on the retail prices of specific goods like washing machines.

One potential puzzle, however, is that some cohort-design studies, such as the Flaaen, Hortaçsu, and Tintelnot (2020) analysis of washing machines, also find near-complete passthrough. This does not necessarily contradict the selection mechanism but suggests its scope may be product- or industry-dependent. There are several potential reasons. First, the firms in that specific market may not have had an incentive to transship in the first place. If the washing machine industry is characterized by a distribution of marginal costs that lies mostly above the avoidance cutoff, then few firms would select into avoidance regardless of the research design. Second, for certain goods, the fixed cost of avoidance may be prohibitively high; large, durable goods like washing machines are difficult to ship discreetly and may require significant, specialized investment to reroute. Finally, passthrough for durable goods, which have complex pricing and supply chains, may be governed by different dynamics than the broader basket of goods driving the aggregate result.

On the other hand, border-level regressions that use customs unit values or aggregate import price indices for targeted product–origin cells estimate β_{border} by construction, since transshippers drop out or are reweighted when the tariff induces origin switching. This is the estimand in the border specifications of Amiti, Redding, and Weinstein (2019), Fajgelbaum et al. (2020), and the unit-value components of Cavallo et al. (2021).

What about exchange rates? The selection logic may also help reconcile the puzzle of why measured tariff passthrough is near unity, while exchange rate passthrough is considerably lower (Cavallo et al. 2021; Gertler 2025). The key difference is the nature of the shock. Tariffs are persistent, so it is worthwhile for firms to invest in costly avoidance infrastructure, such as new transshipment routes. In contrast, exchange rate movements are often transitory and affect all origins, making such large, fixed investments in avoidance unlikely to pay off. The empirical importance these adjustment costs is substantial;

recent evidence from Boehm, Levchenko, and Pandalai-Nayar (2023) shows that the full response of trade to tariff shocks can take nearly a decade to materialize. Therefore, the low passthrough firms pruned from the tariff sample may remain present in the exchange rate sample, explaining why measured passthrough is high in the former and low in the latter.

3.2 Quantifying the Bias

By the law of total expectation,

$$\beta_{\text{cohort}} = (1 - \theta) E[\beta_i \mid S_i = 1] + \theta E[\beta_i \mid S_i = 0], \quad \theta \equiv \Pr(S_i = 0).$$

Using $E[\beta_i \mid S_i = 1] \approx \beta_{\text{border}}$ gives

$$(1 - \theta) \beta_{\text{border}} + \theta \underline{\beta} \leq \beta_{\text{cohort}} \leq \beta_{\text{border}},$$

for any $\underline{\beta} \leq E[\beta_i \mid S_i = 0]$. Under the conservative case $\underline{\beta} = 0$,

$$\beta_{\text{cohort}} \geq (1 - \theta) \beta_{\text{border}}.$$

Implementing this inequality requires only θ and a choice of $\underline{\beta}$. We adopt the conservative case $\underline{\beta} = 0$, so the cohort object is bounded below by $(1 - \theta) \beta_{\text{border}}$. Because we do not observe a single θ , we report a conservative floor by plugging in two transshipment screens: a lower-bound share θ_{low} and a more permissive upper estimate θ_{high} . Using only $E[\beta_i \mid S_i = 0] \geq 0$, the cohort object satisfies

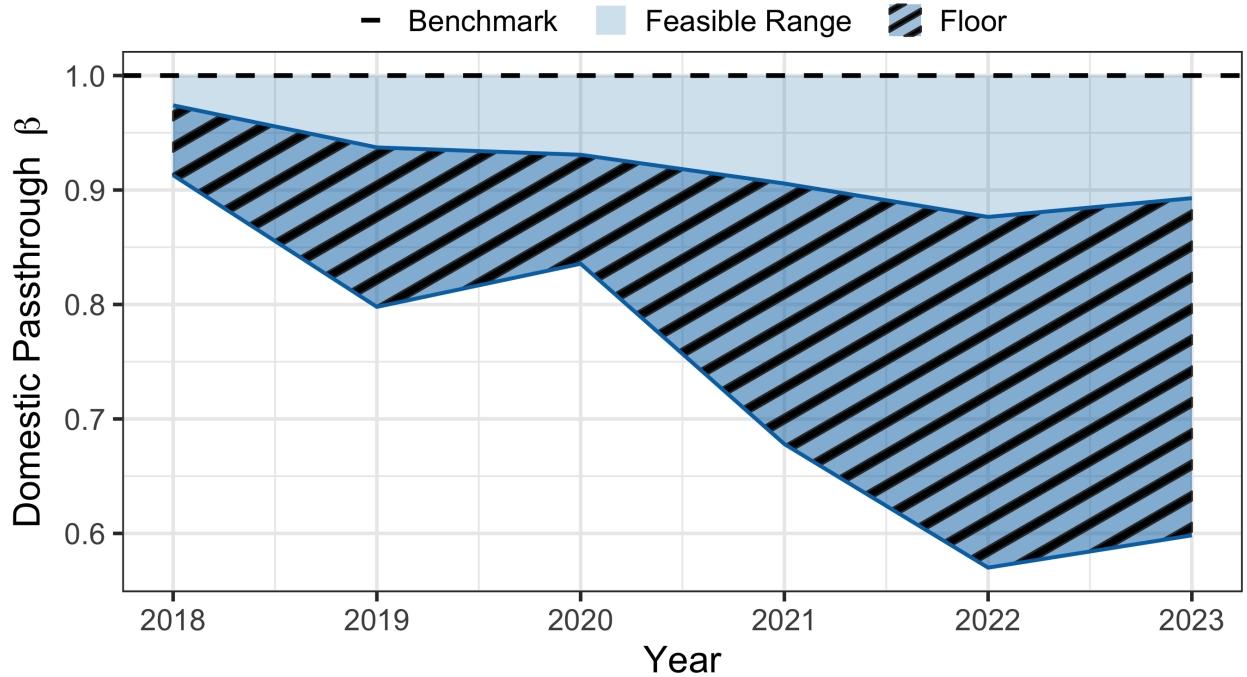
$$\beta_{\text{cohort}} \in [(1 - \theta_{\text{high}}) \beta_{\text{border}}, \beta_{\text{border}}].$$

When we set $\beta_{\text{border}} = 1$, the feasible range is $[1 - \theta_{\text{high}}, 1]$. The interval $[1 - \theta_{\text{high}}, 1 - \theta_{\text{low}}]$ that we shade in Figure 3 should therefore be interpreted as a floor band, not the full feasible range.

To construct the bounds, I follow a methodology from Freund (2024) and use data on trade flows from the U.N. Comtrade database from 2017-2023. I restrict attention to HS6 products where China's tariff rose between December 2017 and December 2018, the "tariffed set." For each year $y \geq 2018$, we then test whether partner exporters exhibit the behavioral patterns expected of transshipment. These follow both a liberal screen, which constitutes an upper bound on transshipment, as well as a conservative screen, which is a lower bound.

The liberal screen follows Freund’s liberal definition. Within each tariffed HS6, I flag partner–product pairs where (i) China’s U.S. import share fell between 2017 and year y while the partner’s share rose, and (ii) China’s global import share rose more than the partner’s. Additionally, it must be that both legs of the potential transshipment route expand relative to 2017: partner exports to the United States must increase, and Chinese exports to the partner must also increase. For flagged pairs, we compute the transshipped value as the minimum of Chinese exports to the partner and the partner’s exports to the United States in year y . Summing across pairs gives an upper bound on the value of transshipped goods. Our conservative screen imposes two additional restrictions. First, we limit partners to a hub whitelist.⁵ Second, we require that Chinese exports to the partner in year y be at least as large as the partner’s exports to the United States. Under this condition the minimum operator is redundant, so the conservative numerator equals the partner’s exports to the United States for flagged pairs. See Freund (2024) for extra details on both construction and on the conditions for transshipment.

Figure 3: Corrected domestic pass-through bounds, 2018–2023



Note: Shaded regions show corrected domestic pass-through relative to the border benchmark. The light blue area indicates the full feasible range implied by the model, $\beta_{\text{cohort}} \in [(1 - \theta_{\text{high}})\beta_{\text{border}}, \beta_{\text{border}}]$. The hatched band marks the conservative floor using our two transshipment screens, $[(1 - \theta_{\text{high}})\beta_{\text{border}}, (1 - \theta_{\text{low}})\beta_{\text{border}}]$. The dashed line shows the border benchmark $\beta_{\text{border}} = 1$.

5. These countries are Vietnam, Mexico, Singapore, Malaysia, Thailand, India, Turkey, United Arab Emirates, Indonesia, Philippines, Israel, Canada, Dominican Republic, Cambodia, Sri Lanka, and Taiwan.

Corrected pass-through is roughly 90–95% in 2018, then falls. Tariffs started mid-year and avoidance takes time, so Figure 3 should be read as adjustment toward a long-run level. The liberal screen likely overcounts avoidance by picking up relocation, third-country FDI, or supplier switching, so the truth is closer to θ_{low} . The key object is the band: even the conservative screen shows transshipment above 10% by 2019, trimming measured pass-through by 10–15 pp relative to the border estimate.

To validate the methodology, I perform a placebo test by applying the same screens to a set of untreated goods. The results are in Appendix Figure A.1. The methodology identifies an economically insignificant amount of transshipment for these products, with the liberal methodology only falsely identifying more than 5% of goods by 2022. This null result for the placebo group provides strong evidence that the main estimates are capturing a real, policy-induced behavioral response and are not an artifact of the screening methodology itself.

The finding that 10–40% of target trade is plausibly rerouted is a first-order empirical issue. However, this large-scale avoidance is not just a measurement problem; it creates a significant welfare problem by eroding the tariff base. This leakage represents a major fiscal externality that increases the tariff’s true welfare cost. We now turn to analyzing these welfare implications and a potential policy response.

4 Enforcement via Tax Deductibility

The welfare cost of this avoidance can be analyzed through the Marginal Cost of Public Funds (MCPF), which combines the domestic incidence of the tax (the numerator) and the fiscal externality from the behavioral response (the denominator). The previous section’s findings directly correct our measurement of the numerator. This section now turns to the denominator, analyzing the fiscal externality itself and a novel enforcement margin that can reduce it.

4.1 Welfare and the Fiscal Externality

Finkelstein and Hendren (2020) define the marginal cost of public funds of any given tax as

$$\text{MCPF} = \frac{\gamma}{1 + \eta},$$

where γ is the domestic incidence of the tax and $\eta < 0$ is the elasticity of the tariff base (the value of imports subject to the duty).⁶ The MCPF is the ratio of the marginal welfare cost of a tax to the marginal benefit of revenue raised from that instrument. To first order, the numerator is simply the domestic tax incidence. Thus, for a tariff, $\gamma < 1$. In fact, in our context, we have shown that $\gamma = \beta_{\text{cohort}}$. That is important because standard estimates of the welfare cost of tariffs will typically overrate how costly they are on the margin if the estimates are based on upwardly biased estimates of passthrough. Consequently, the estimates of the MCPF from Finkelstein and Hendren (2020) ought to be scaled down by the range $[1 - \theta_{\text{high}}, 1 - \theta_{\text{low}}]$. Since they estimate a tariff MCPF of around 1.2-1.6, a new range would be from 0.7-1.6.⁷ Note that if the MCPF is less than one, then it is cheaper in welfare terms than a lump-sum tax, which has an MCPF of one. However, the MCPF of a tariff is still likely far above one.

However, the 0.7-1.6 range holds fixed the fiscal externality. What if it could be reduced via policy and hence reduce the welfare cost? With a tariff, there are two components to the fiscal externality. The first is simply replacing targeted goods with untargeted foreign or domestic competitors. The second, which is relevant to us, is avoidance behavior like transshipment. In [Executive Order 14326 \(2025\)](#), the government tried to minimize the fiscal externality by issuing a levy on identified transshippers. That strategy is aligned with the tax avoidance literature (Slemrod and Yitzhaki 2002). However, the tax system may already interact with tariffs such that transshipment is largely minimized. I turn to analyzing that subsequently.

4.2 Tax Deductibility as an Enforcement Tool

The routing decision turns on the gap between the direct channel, where buyers face the tariff wedge, and the avoidance benchmark, where the wedge is lower but a fixed cost F is paid. Policies that compress this gap reduce transshipment and hence raise tariff revenue. Canonically, the public finance literature would advocate for penalties or nudges to reduce avoidance (Slemrod and Yitzhaki 2002), but the existing tax system may unintentionally do much of the work already. In this subsection, I emphasize that interaction between domestic business policies like bonus depreciation and tariffs reduces transshipment via avoidance.

In the spirit of compressing the gap via penalties, let the direct channel carry an ad

6. They use the terminology marginal *value* of public funds. I use “cost” here for the sake of clarity, but the concept is the same. See Appendix B for a derivation.

7. That is smaller than the range of estimates for top marginal tax rates, and considerably smaller than those for both corporate taxes and capital gains taxes.

valorem wedge $t_D \geq 0$ and the transshipment channel a (weakly) smaller wedge $t_T \in [0, t_D)$. Delivered prices are $(1 + t_D)p$ under direct and $(1 + t_T)p$ under avoidance.

For any marginal cost m , define the optimized value in a channel with delivered-price multiplier $1 + t$:

$$V(m, 1 + t) \equiv \max_{p \geq m} (p - m) D((1 + t)p), \quad \pi_D(m) = V(m, 1 + t_D), \quad \pi_T(m) = V(m, 1 + t_T) - F,$$

and the avoid–direct difference

$$\Delta(m; t_D, t_T) \equiv \pi_T(m) - \pi_D(m) = V(m, 1 + t_T) - V(m, 1 + t_D) - F.$$

Proposition 4 (Cutoff routing and wedge comparative statics). *Under D1–D4, there exists a unique cutoff $\hat{m}(t_D, t_T) \in [m, \bar{m}]$ such that*

$$m < \hat{m}(t_D, t_T) \Rightarrow \text{Transship}, \quad m \geq \hat{m}(t_D, t_T) \Rightarrow \text{Direct}.$$

If the cutoff is interior, then

$$\frac{\partial \hat{m}}{\partial t_D} > 0 \quad \text{and} \quad \frac{\partial \hat{m}}{\partial t_T} < 0.$$

Proof: See Appendix C.4.

Let $\theta(t_D, t_T) \equiv \Pr(m < \hat{m}(t_D, t_T))$. By the signs in Proposition 4, $\partial\theta/\partial t_D > 0$ and $\partial\theta/\partial t_T < 0$: compressing the direct–transship wedge reduces transshipment mechanically.

Deductibility as raising the transshipment wedge. Suppose the buyer is a domestic firm facing corporate rate $\tau_c \in [0, 1)$ with partial deductibility $z \in [0, 1]$ of tariff-inclusive input costs. Here $z \in [0, 1]$ is the NPV factor on deductions (with $z = 1$ under full expensing). The factor $(1 - z\tau_c)$ multiplies all input costs, whether or not the tariff is paid; the difference arises only from the tariff wedge t_D . After-tax unit costs are

$$C_D = (1 - \tau_c z)(1 + t_D)p, \quad C_T = (1 - \tau_c z)p,$$

so the direct–avoid cost gap under deductibility is

$$\Delta C_{\text{ded}} = C_D - C_T = (1 - \tau_c z) t_D p.$$

In the two-wedge model where buyers face delivered prices $(1 + t)p$ directly, the corresponding gap is

$$\Delta C_{\text{wedge}} = [(1 + t_D) - (1 + t_T)]p = (t_D - t_T)p.$$

Equating gaps yields a simple isomorphism.

Corollary 1 (Deductibility–wedge isomorphism). *With full deductibility, the routing problem is behaviorally equivalent to a two-wedge problem with*

$$t_T = z\tau_c t_D,$$

and with no deductibility the benchmark is $t_T = 0$.

Combining Corollary 1 with Proposition 4 turns the corporate-tax knob into a clean comparative static:

$$\frac{d\hat{m}}{d\tau_c} < 0 \quad \text{and} \quad \frac{d\hat{m}}{dz} < 0.$$

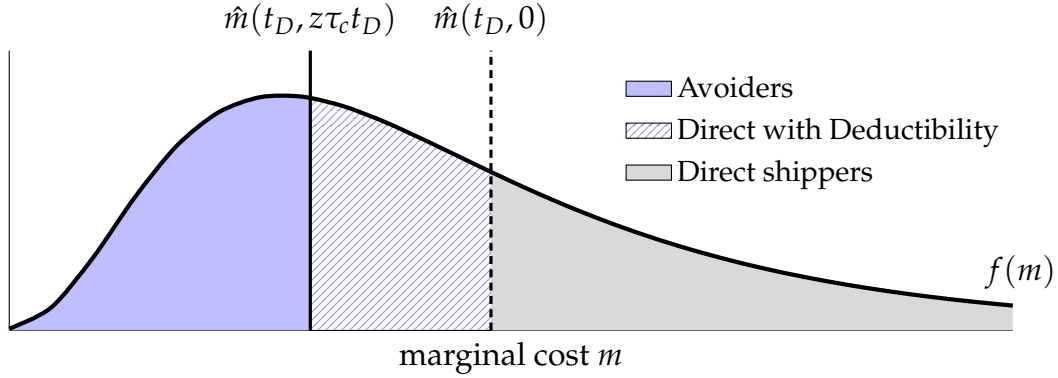
A higher tax rate ($\tau_c \uparrow$) or a higher deductibility fraction z reduces transshipment: the cutoff moves left, the survivor set expands, and the marginal transshipper is less inclined to re-route because deductibility effectively loads a share $z\tau_c$ of the tariff into the avoidance channel. That is easy to see in Figure 4, where the cutoff for transshipment shifts left when intermediates are tax-deductible. However, the effect only applies in practice to intermediate goods and imported equipment; households do not have a feasible path to rebate tariffs.⁸ When deductibility rises, as with bonus depreciation, then transshipment declines.

A decline in θ also has implications for revenue and welfare. Volume moves from avoidance to compliant shipments, so customs receipts on the targeted flow rise, while the corporate base narrows mechanically through expensing. The net fiscal effect is ambiguous ex ante and depends on the prevalence of deductible inputs in the tariffed set and on τ_c . For welfare accounting, $\gamma = \beta_{\text{cohort}}$ is a property of the pre-tariff cohort and does not vary with θ . What deductibility changes is the fiscal externality by reducing avoidance (lower θ). A lower θ also improves measurement because border estimates move closer to β_{cohort} . Deductibility therefore acts like an enforcement tool that reduces transshipment, aligns border measures with the cohort object, tilts revenue toward customs,

8. Importantly, the business tax penalty on transshipment is monotonically increasing in the corporate tax rate. Suppose there is full expensing. Abel (2007) shows that this paired with any corporate tax rate is optimal in a closed economy. It could plausibly be the case that with tariffs, a positive corporate tax rate with full expensing is optimal.

and lowers the marginal welfare cost of tariff revenue.

Figure 4: An increase in tax-deductibility shifts the cutoff left



Note: Distribution of marginal costs and tariff-induced selection. The cutoff shifts left along the marginal cost distribution when tariffs are tax-deductible.

Corollary 1 has direct policy content and is directly applicable to our last two major tax reforms. The 2017 Tax Cuts and Jobs Act's 100% bonus depreciation made many imported equipment purchases immediately deductible, reducing the payoff to transshipment. Importantly, the One Big Beautiful Bill Act made that change permanent.⁹

5 Conclusion

This paper shows that tariff avoidance has first-order implications for both measurement and policy. Selection into transshipment biases border estimates of tariff passthrough upward, and bounding the resulting correction reveals that true incidence is materially lower than the literature suggests. The same behavior erodes the tariff base and raises the fiscal externality, which the domestic tax system can partially offset by making tariff-inclusive input costs deductible. While the analysis here is static and partial-equilibrium, future work should examine how the persistence of tariffs shapes firms' dynamic investment in avoidance and how large-scale transshipment affects wages, prices, and trade flows in a general equilibrium setting. These extensions would enrich the welfare analysis, but the present framework is sufficient to establish the core selection mechanism and its first-order implications for measurement and policy.

9. Alessandria et al. (2025) emphasize the efficiency of pairing tariffs with investment subsidies in the spirit of Diamond and Mirrlees (1971); the isomorphism here shows expensing itself acts on the avoidance margin.

References

- Abel, Andrew. 2007. *Optimal Capital Income Taxation*. Technical report. Cambridge, MA: National Bureau of Economic Research, August. <https://doi.org/10.3386/w13354>.
- Alessandria, George, Jiaxiaomei Ding, Shafaat Khan, and Carter Mix. 2025. *The Tariff Tax Cut: Tariffs as Revenue*. Technical report. Cambridge, MA: National Bureau of Economic Research, May. <https://doi.org/10.3386/w33784>.
- Allingham, Michael G., and Agnar Sandmo. 1972. "Income tax evasion: a theoretical analysis." *Journal of Public Economics* 1, nos. 3-4 (November): 323–338. ISSN: 00472727. [https://doi.org/10.1016/0047-2727\(72\)90010-2](https://doi.org/10.1016/0047-2727(72)90010-2).
- Amiti, Mary, Oleg Itskhoki, and Jozef Konings. 2019. "International Shocks, Variable Markups, and Domestic Prices." *The Review of Economic Studies* 86, no. 6 (November): 2356–2402. ISSN: 0034-6527. <https://doi.org/10.1093/restud/rdz005>.
- Amiti, Mary, Stephen J. Redding, and David E. Weinstein. 2019. "The Impact of the 2018 Tariffs on Prices and Welfare." *Journal of Economic Perspectives* 33, no. 4 (November): 187–210. ISSN: 0895-3309. <https://doi.org/10.1257/jep.33.4.187>.
- Bhandari, Anmol, David Evans, Ellen R Mcgrattan, and Yuki Yao. 2024. *Business Income Underreporting and Public Finance**. Technical report.
- Boehm, Christoph E., Andrei A. Levchenko, and Nitya Pandalai-Nayar. 2023. "The Long and Short (Run) of Trade Elasticities." *American Economic Review* 113, no. 4 (April): 861–905. ISSN: 0002-8282. <https://doi.org/10.1257/aer.20210225>.
- Cavallo, Alberto, Gita Gopinath, Brent Neiman, and Jenny Tang. 2021. "Tariff Pass-Through at the Border and at the Store: Evidence from US Trade Policy." *American Economic Review: Insights* 3, no. 1 (March): 19–34. ISSN: 2640-205X. <https://doi.org/10.1257/aeri.20190536>.
- De Loecker, Jan, Jan Eeckhout, and Gabriel Unger. 2020. "The Rise of Market Power and the Macroeconomic Implications*." *The Quarterly Journal of Economics* 135, no. 2 (May): 561–644. ISSN: 0033-5533. <https://doi.org/10.1093/qje/qjz041>.
- Diamond, Peter A., and James A. Mirrlees. 1971. "Optimal Taxation and Public Production I: Production Efficiency." *American Economic Review* 61 (1): 8–27.
- Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu. 2023. "How Costly Are Markups?" *Journal of Political Economy* 131, no. 7 (July): 1619–1675. ISSN: 0022-3808. <https://doi.org/10.1086/722986>.
- Executive Order 14326. 2025. *Further Modifying the Reciprocal Tariff Rates*, July.
- Fajgelbaum, Pablo, and Amit Khandelwal. 2024. *The Value of De Minimis Imports*. Technical report. Cambridge, MA: National Bureau of Economic Research, June. <https://doi.org/10.3386/w32607>.
- Fajgelbaum, Pablo D, Pinelopi K Goldberg, Patrick J Kennedy, and Amit K Khandelwal. 2020. "The Return to Protectionism*." *The Quarterly Journal of Economics* 135, no. 1 (February): 1–55. ISSN: 0033-5533. <https://doi.org/10.1093/qje/qjz036>.
- Finkelstein, Amy, and Nathaniel Hendren. 2020. "Welfare Analysis Meets Causal Inference." *Journal of Economic Perspectives* 34, no. 4 (November): 146–167. ISSN: 0895-3309. <https://doi.org/10.1257/jep.34.4.146>.
- Fisman, Raymond, and Shang-Jin Wei. 2004. "Tax Rates and Tax Evasion: Evidence from "Missing Imports" in China." *Journal of Political Economy* 112, no. 2 (April): 471–496. ISSN: 0022-3808. <https://doi.org/10.1086/381476>.

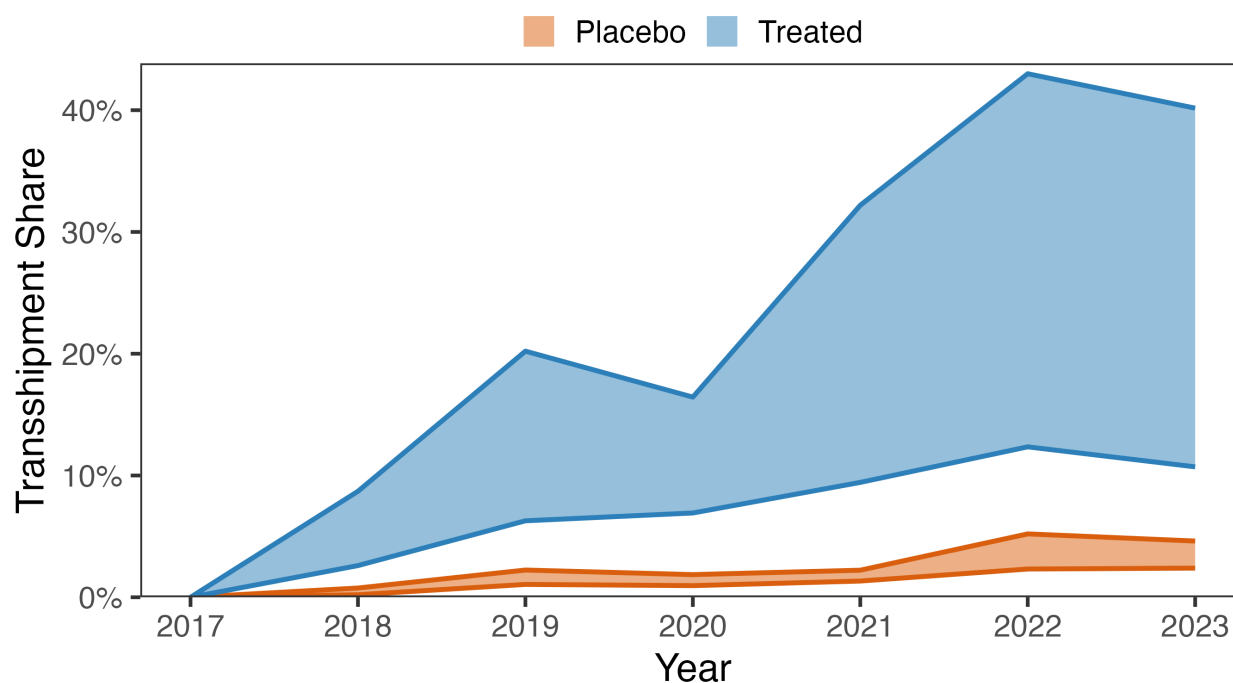
- Flaaen, Aaron, Ali Hortaçsu, and Felix Tintelnot. 2020. "The Production Relocation and Price Effects of US Trade Policy: The Case of Washing Machines." *American Economic Review* 110, no. 7 (July): 2103–2127. ISSN: 0002-8282. <https://doi.org/10.1257/aer.20190611>.
- Freund, Caroline. 2024. *The China Wash: Tracking Products To Identify Tariff Evasion Through Transshipment*. Technical report.
- Gertler, Sarah M. 2025. "Exchange Rate Pass-through and Expenditure-Switching Revisited."
- Gopinath, Gita, and Oleg Itskhoki. 2010. "Frequency of Price Adjustment and Pass-Through." *Quarterly Journal of Economics* 125, no. 2 (May): 675–727. ISSN: 0033-5533. <https://doi.org/10.1162/qjec.2010.125.2.675>.
- Hendren, Nathaniel. 2016. "The Policy Elasticity." *Tax Policy and the Economy* 30, no. 1 (January): 51–89. ISSN: 0892-8649. <https://doi.org/10.1086/685593>.
- Iyoha, Ebehi, Edmund Malesky, Jaya Wen, and Sung-Ju Wu. 2025. *Exports in Disguise? Trade Rerouting during the US-China Trade War*. Technical report.
- Kimball, Miles S. 1995. "The Quantitative Analytics of the Basic Neomonetarist Model." *Journal of Money, Credit and Banking* 27, no. 4 (November): 1241. ISSN: 00222879. <https://doi.org/10.2307/2078048>.
- Slemrod, Joel. 2019. "Tax Compliance and Enforcement." *Journal of Economic Literature* 57, no. 4 (December): 904–954. ISSN: 0022-0515. <https://doi.org/10.1257/jel.20181437>.
- Slemrod, Joel, and Shlomo Yitzhaki. 2002. "Tax Avoidance, Evasion, and Administration," 1423–1470. [https://doi.org/10.1016/S1573-4420\(02\)80026-X](https://doi.org/10.1016/S1573-4420(02)80026-X).

A Placebo Test on Untreated Goods

As a falsification exercise, we apply our transshipment screens to HS6 products that were never subject to Section 301 tariff increases. We define the placebo set as the complement of the tariffed list—all HS6 where China’s tariff did *not* rise between December 2017 and December 2018. For this untreated set, we compute transshipment bands using the same procedures as for the treated set: (i) U.S. import share shifts (China’s share down, partner’s share up), (ii) global share dominance (China’s share rising more than the partner’s), and (iii) the liberal or conservative flow restrictions. The numerator is the sum of flagged flows, and the denominator is China’s direct exports to the U.S. in 2017 on the untreated HS6 set, deflated to 2017 dollars.

Figure A.1 shows that the placebo transshipment shares remain close to zero throughout 2018–2023, with narrow bands under both the conservative and liberal definitions. This contrasts with the treated HS6 set, which displays steadily rising transshipment bands following the tariff shocks. The placebo test confirms that our screening procedure does not spuriously generate transshipment where no tariff-induced transshipment margin exists.

Figure A.1: Transshipment share with placebo test, 2018–2023



Note: Placebo transshipment bands for untreated HS6 products, 2017–2023. The placebo set is defined as the complement of the tariffed list—HS6 categories where China’s tariff did not rise between December 2017 and December 2018.

B The Marginal Cost of Public Funds

Finkelstein and Hendren (2020) and Hendren (2016) give a thorough overview of the marginal *value* of public funds. Because I focus on taxes rather than transfers, I call it the marginal *cost* for the sake of clarity. Here, I briefly derive the MCPF formula in the main text; see the aforementioned sources for further details. Suppose a planner has access to a vector of linear tax instruments τ (with individual elements τ_i). Each tax instrument has a corresponding base $B_i(\tau)$. For simplicity, I assume no spillovers between them. The government chooses instruments to maximize a well-behaved welfare function $W(\tau)$ subject to a revenue constraint:

$$\max_{\tau} W(\tau) \quad \text{subject to} \quad G = R(\tau) \equiv \sum_{i=1}^N \tau_i B_i(\tau) \quad (\text{SPP})$$

Denote λ as the shadow value of public funds:

$$\lambda = -\frac{W'(\tau_i)}{R'(\tau_i)} \iff \frac{W'(\tau_i)}{R'(\tau_i)} = \frac{W'(\tau_j)}{R'(\tau_j)}.$$

With no spillovers,

$$R'(\tau_i) = B_i(\tau_i) + \tau_i \frac{\partial B_i}{\partial \tau_i} = B_i(\tau_i) \times \underbrace{\left(1 + \tau_i \frac{B'_i(\tau_i)}{B_i(\tau_i)}\right)}_{1+\eta_i}$$

By the envelope theorem, the marginal welfare cost of nudging each instrument is the domestic willingness to pay to avoid it, i.e., $-\partial W / \partial \tau_i = \gamma_i B_i$, so

$$\text{MCPF}_{\tau_i} = \frac{\gamma_i B_i(\tau_i)}{B_i(\tau_i)(1 + \eta_i)} = \frac{\gamma_i}{1 + \eta_i}.$$

C Proofs

C.1 Proof of Lemma 1.

Proof. Interior and FOC. By **D2–D4**, $\Phi_i(\cdot)$ is strictly concave in p_i ; the unique maximizer is interior and characterized by the FOC. Let the delivered price be

$$\tilde{p}_i \equiv \tilde{p}_i(p_i, \tau) = \begin{cases} p_i & \text{(transship channel)} \\ (1 + \tau) p_i & \text{(direct channel)}. \end{cases}$$

Then $\partial \tilde{p}_i / \partial p_i \in \{1, 1 + \tau\}$, and the FOC is

$$0 = \frac{\partial \Phi_i}{\partial p_i} = D_i(\tilde{p}_i) + (p_i - m_i) D'_i(\tilde{p}_i) \frac{\partial \tilde{p}_i}{\partial p_i}.$$

With $\varepsilon_i(\tilde{p}_i) \equiv -\tilde{p}_i D'_i(\tilde{p}_i) / D_i(\tilde{p}_i)$,

$$\frac{p_i^* - m_i}{p_i^*} = \frac{\tilde{p}_i}{(\partial \tilde{p}_i / \partial p_i) p_i^*} \cdot \frac{-D_i(\tilde{p}_i)}{\tilde{p}_i D'_i(\tilde{p}_i)} = \frac{1}{\varepsilon_i(\tilde{p}_i)}.$$

This is the Lerner condition.

Monotonicity in m_i . Define

$$g(p, m; \tau) \equiv D(\tilde{p}) + (p - m) D'(\tilde{p}) \frac{\partial \tilde{p}}{\partial p}, \quad \tilde{p} = \tilde{p}(p, \tau).$$

At the optimum, $g(p^*, m; \tau) = 0$. Partial derivatives:

$$\frac{\partial g}{\partial m} = -D'(\tilde{p}) \frac{\partial \tilde{p}}{\partial p} > 0 \quad (\mathbf{D1}), \quad \frac{\partial g}{\partial p} = 2 D'(\tilde{p}) \frac{\partial \tilde{p}}{\partial p} + (p - m) D''(\tilde{p}) \left(\frac{\partial \tilde{p}}{\partial p} \right)^2.$$

Strict concavity of profits (**D4**) implies $\partial g / \partial p < 0$ at p^* . By the IFT, $\partial p^* / \partial m = -(\partial g / \partial m) / (\partial g / \partial p) > 0$.

Monotonicity in τ (direct channel). For the direct channel, $\tilde{p} = (1 + \tau)p$. Write

$$G(p, \tau) \equiv \frac{p - m}{p} - \frac{1}{\varepsilon((1 + \tau)p)} = 0.$$

Then

$$\frac{\partial G}{\partial \tau} = \frac{\varepsilon'((1 + \tau)p)}{\varepsilon((1 + \tau)p)^2} p \geq 0, \quad \frac{\partial G}{\partial p} = \frac{m}{p^2} + \frac{\varepsilon'((1 + \tau)p)}{\varepsilon((1 + \tau)p)^2} (1 + \tau).$$

Under Kimball demand, $\varepsilon'(\cdot) \geq 0$, so $\partial G / \partial \tau \geq 0$. Profit concavity (**D4**, equivalently

$\varepsilon - 1 - \kappa > 0$ at the optimum) implies $\partial G / \partial p > 0$. Hence, by the IFT,

$$\frac{\partial p^*}{\partial \tau} = - \frac{\partial G / \partial \tau}{\partial G / \partial p} \leq 0,$$

with equality under CES ($\varepsilon' \equiv 0$).

Quantities. In the direct channel $\tilde{p} = (1 + \tau)p^*$ and

$$\frac{d((1 + \tau)p^*)}{d\tau} = p^* + (1 + \tau) \frac{dp^*}{d\tau} = p^* \left(1 + \frac{\kappa(\tilde{p})}{\varepsilon(\tilde{p}) - 1 - \kappa(\tilde{p})} \right) = p^* \frac{\varepsilon(\tilde{p}) - 1}{\varepsilon(\tilde{p}) - 1 - \kappa(\tilde{p})} > 0,$$

since $\varepsilon(\tilde{p}) > 1$ and the denominator is positive by D4. With $D'(\cdot) < 0$, $q_i^* = D(\tilde{p})$ strictly falls in τ ; similarly, q_i^* strictly falls in m because p^* rises in m and demand slopes down. This proves (i)–(iii). \square

C.2 Proof of Proposition 1

Proof. Define the value function in the direct channel

$$V(m, 1 + \tau) \equiv \max_{p \geq m} (p - m) D((1 + \tau)p),$$

and set $\pi^D(m) = V(m, 1 + \tau)$, $\pi^T(m) = V(m, 1) - F$, $\Delta(m; \tau) = \pi^T(m) - \pi^D(m)$. By **D3–D4**, for every (m, τ) the maximizer $p^*(m, \tau)$ is interior and unique; hence V is continuously differentiable in $(m, 1 + \tau)$ and the envelope theorem applies.

Increasing differences. For the primitive $(p - m)D((1 + \tau)p)$ with choice p and parameters $(m, 1 + \tau)$,

$$\frac{\partial^2}{\partial m \partial (1 + \tau)} \left[(p - m)D((1 + \tau)p) \right] = -p D'((1 + \tau)p) > 0 \quad (\mathbf{D1}),$$

so the primitive has increasing differences in $(m, 1 + \tau)$. By Topkis, $V(m, 1 + \tau)$ inherits increasing differences.

(a) Fix τ . Using the envelope theorem,

$$\frac{\partial \Delta}{\partial m} = V_m(m, 1) - V_m(m, 1 + \tau).$$

Since V has increasing differences and $1 < 1 + \tau$, it follows that $V_m(m, 1) < V_m(m, 1 + \tau)$, so $\partial \Delta / \partial m < 0$. Continuity of V (from **D2–D4**) implies continuity of $\Delta(\cdot; \tau)$, hence $\{m :$

$\Delta(m; \tau) \geq 0\}$ is an interval. Define

$$\hat{m}(\tau) \equiv \sup\{m : \Delta(m; \tau) > 0\} \in [\underline{m}, \bar{m}].$$

Strict monotonicity of $\Delta(\cdot; \tau)$ delivers uniqueness of the cutoff and the stated threshold rule.

(b) At the cutoff $\Delta(\hat{m}(\tau); \tau) = 0$. Differentiate implicitly:

$$\frac{d\hat{m}}{d\tau} = - \frac{\partial \Delta / \partial \tau}{\partial \Delta / \partial m}.$$

By the envelope theorem applied at $p^*(m, \tau)$,

$$\frac{\partial V(m, 1 + \tau)}{\partial(1 + \tau)} = (p^* - m) p^* D'((1 + \tau)p^*) \leq 0,$$

so $\partial V(m, 1 + \tau) / \partial \tau \leq 0$, while $\partial V(m, 1) / \partial \tau = 0$. Therefore

$$\frac{\partial \Delta}{\partial \tau} = - \frac{\partial V(m, 1 + \tau)}{\partial \tau} \geq 0.$$

Part (a) gave $\partial \Delta / \partial m < 0$. Hence $d\hat{m} / d\tau > 0$. The monotonicity of S_i and $\theta(\tau)$ follows immediately from the threshold characterization. \square

C.3 Proof of Lemma 2

Lemma 2 (Passthrough under ad valorem tariffs). *Let the delivered price in the direct channel be $\tilde{p}_D \equiv (1 + \tau) p^*(m; D)$. Define $\varepsilon(u) \equiv -u D'(u) / D(u)$ and $\kappa(u) \equiv -\partial \ln \varepsilon(u) / \partial \ln u$. For any survivor,*

$$\beta \equiv \frac{d \ln \tilde{p}_D}{d\tau} = \frac{1}{1 + \tau} \cdot \frac{\varepsilon(\tilde{p}_D) - 1}{\varepsilon(\tilde{p}_D) - 1 - \kappa(\tilde{p}_D)}.$$

Moreover, the denominator is strictly positive at the optimum under D4 (profit concavity), which is equivalent to $\varepsilon(\tilde{p}_D) - 1 - \kappa(\tilde{p}_D) > 0$.

Proof. Write the FOC as $G(p, \tau) \equiv (p - m) / p - 1 / \varepsilon((1 + \tau)p) = 0$. Totally differentiate and use

$$\frac{\partial G}{\partial \tau} = \frac{\varepsilon'((1 + \tau)p)}{\varepsilon((1 + \tau)p)^2} p, \quad \frac{\partial G}{\partial p} = \frac{m}{p^2} + \frac{\varepsilon'((1 + \tau)p)}{\varepsilon((1 + \tau)p)^2} (1 + \tau).$$

Using $\kappa(u) = -u \varepsilon'(u) / \varepsilon(u)$ and the Lerner condition $m = p^*(\varepsilon - 1) / \varepsilon$,

$$\frac{dp^*}{d\tau} = - \frac{\partial G / \partial \tau}{\partial G / \partial p} = \frac{\kappa(\tilde{p}_D)}{1 + \tau} \cdot \frac{p^*}{\varepsilon(\tilde{p}_D) - 1 - \kappa(\tilde{p}_D)}.$$

Finally,

$$\beta = \frac{d \ln((1 + \tau)p^*)}{d\tau} = \frac{1}{1 + \tau} + \frac{1}{p^*} \frac{dp^*}{d\tau} = \frac{1}{1 + \tau} \cdot \frac{\varepsilon(\tilde{p}_D) - 1}{\varepsilon(\tilde{p}_D) - 1 - \kappa(\tilde{p}_D)}.$$

Under D4, strict profit concavity at the optimum is equivalent to $\varepsilon - 1 - \kappa > 0$, which ensures the denominator is positive. \square

C.4 Proof of Proposition 4

Proof. By **D2–D4**, the maximizer in $V(m, 1 + t)$ is interior and unique; envelope arguments apply. As in Proposition 1, the primitive $(p - m)D((1 + t)p)$ has increasing differences in $(m, 1 + t)$ because $\partial^2[(p - m)D((1 + t)p)]/\partial m \partial(1 + t) = -pD'((1 + t)p) > 0$ by **D1**. Hence V inherits increasing differences in $(m, 1 + t)$ by Topkis. Fix (t_D, t_T) . Then

$$\frac{\partial \Delta}{\partial m} = V_m(m, 1 + t_T) - V_m(m, 1 + t_D) < 0$$

since $1 + t_T < 1 + t_D$ and V has increasing differences; continuity of V yields a unique threshold \hat{m} solving $\Delta(\hat{m}; t_D, t_T) = 0$. For the comparative statics, differentiate the indifference condition:

$$\frac{d\hat{m}}{dt_D} = -\frac{\partial \Delta / \partial t_D}{\partial \Delta / \partial m}, \quad \frac{d\hat{m}}{dt_T} = -\frac{\partial \Delta / \partial t_T}{\partial \Delta / \partial m}.$$

By the envelope theorem at the direct and transship optima p_D^* and p_T^* ,

$$\frac{\partial V(m, 1 + t)}{\partial(1 + t)} = (p^* - m) p^* D'((1 + t)p^*) \leq 0,$$

so $\partial V(m, 1 + t_D)/\partial t_D \leq 0$ and $\partial V(m, 1 + t_T)/\partial t_T \leq 0$. Hence

$$\frac{\partial \Delta}{\partial t_D} = -\frac{\partial V(m, 1 + t_D)}{\partial t_D} \geq 0, \quad \frac{\partial \Delta}{\partial t_T} = \frac{\partial V(m, 1 + t_T)}{\partial t_T} \leq 0.$$

Since $\partial \Delta / \partial m < 0$ at an interior cutoff, the stated signs follow. \square