Selection into Transshipment 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 that this finding may be partly due to selection bias. When exporters can reroute goods through a third country—a practice known as "transshipment"—the third country is recorded as the origin, and both the exporter and importer avoid the tariff. With heterogeneous markups, exporters bearing the most incidence are the most likely to transship. Because these firms also exhibit the lowest passthrough, their exit from the direct-shipping sample biases border estimates upward, making it appear that domestic consumers bear more of the tariff than they actually do. Using data from Comtrade and a methodology for identifying transshipment developed by Freund (2024), I estimate that accounting for transshipment reduces the measured tariff incidence on U.S. importers from 100% to between 60% and 90%. This selection logic points to an enforcement margin: policies that narrow the cost gap between direct shipping and avoidance reduce transshipment. I extend the model to show that allowing full tax deductibility of imported intermediate and investment goods achieves exactly that. This raises compliant flows and lowers the measured welfare cost of tariffs on the margin.

*Massachusetts Institute of Technology, jpmejia@mit.edu. First Draft: September 2025. I am grateful to Steve Miran and Kim Ruhl for prompting initial discussions that led to 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

Foreign exporters may take a variety of creative steps to avoid tariffs. For example, Fisman and Wei (2004) show that exporters may misreport the value of goods, or misclassify them, to get around tariffs. More recently, foreign exporters have begun shipping goods through other countries with lower tariff rates than their own. This practice, known as *transshipping*, 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). That practice has implications for measuring tariff incidence: if low-markup firms choose to transship, then the measured passthrough of tariffs into domestic prices is biased upward. That matters because typical estimates of tariff passthrough place it around 100%, meaning that domestic consumers and producers bear the full incidence of tariffs and foreign exporters bear none (Amiti, Redding, and Weinstein 2019; Cavallo et al. 2021; Fajgelbaum et al. 2020).

I propose a simple framework in which firms are monopolistically competitive with heterogeneous marginal costs, and they make two choices. First, they calculate optimal prices that maximize profits under two regimes. The first regime they consider is the one in which they ship directly and pay a tariff. Under the second is one where they pay a fixed cost to avoid the tariff by shipping their product through a third country. Then, the firm compares profits under the two scenarios. Under standard regularity assumptions, some share of firms will choose to transship. The key question is whether markups are heterogeneous and if that results in selection into transshipping. If markups are heterogeneous, like with Kimball (1995) demand, then firms with low markups will choose to transship, while firms with high markups will not. The intuition is straightforward: high-markup firms exhibit low pass-through. When a tariff raises the delivered price and demand becomes more elastic, they trim markups and absorb more of the tax, while higher-cost, lower-markup firms pass through a larger share to domestic buyers. As a result, exporters which bear the brunt of any tariff choose to transship and disappear from most datasets.

Heterogeneous markups are a sufficient condition for bias in empirical pass-through estimates whenever exporters can choose to transship and the data only observe direct shipments at the border.[1] The bias is design-specific, so we must distinguish *cohort-*

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. However, it is far from obvious what the sign of heterogeneous markups is for welfare (Albrecht, Phelan, and Pretnar 2023). In fact, the mechanism in this paper suggests that heterogeneous foreign markups are beneficial for domestic welfare.

from *border*-level regressions. Cohort designs track the same exporter–importer (or product–firm) unit across routing so the item remains in sample whether it ships direct or via a hub; in that case, selection into transshipment does not distort the estimand, as in parts of Cavallo et al. (2021) and Flaaen, Hortaçsu, and Tintelnot (2020). By contrast, border designs—which include Amiti, Redding, and Weinstein (2019), Fajgelbaum et al. (2020), and components of Cavallo et al. (2021)—rely on customs unit values or aggregate import price indices constructed from direct shipments. Because low-markup, low–passthrough exporters are the ones most likely to re-route, the estimating sample is composed of survivors and the border coefficient overstates the cohort object. This same selection logic helps reconcile the long-standing puzzle of low exchange-rate pass-through and high tariff pass-through: exchange-rate movements rarely induce rerouting infrastructure, while tariffs do, pruning exactly the observations that would have muted pass-through.

That raises the question of how large transshipping is, and consequently, how much smaller the incidence must be on domestic importers than what we measure in the data. Using a methodology developed by Freund (2024), I show that between 10-40% of goods targeted by the 2018 tariffs on China were transshipped or eventually transshipped. Theory provides a sufficient statistic approach for how to translate that range into a correction on measured passthrough. If we have biased estimates of passthrough, then we can provide bounds on how biased it is by simply dividing the value of transshipped goods by the pre-tariff export share from China, which is exactly what the 10-40% range is. That implies actual passthrough is between 60-90%, rather than the 100% documented by a large literature on the incidence of the 2018 Chinese tariffs.

Correcting upwardly biased tariff incidence estimates has important welfare implications. Typically, we evaluate the welfare cost of tax instruments by considering their corresponding marginal cost of public funds (MCPF) (Finkelstein and Hendren 2020). The MCPF combines two elements: the domestic incidence of taxation, and the fiscal externality of raising revenue. In principle, the higher domestic incidence, the higher the welfare cost of tariffs. Simply applying the correction implies that the MCPF of tariff revenue is about 60-90% as large as commonly understood, which may make it a cheaper tool to use on the margin than many other types of tax instruments. The other component is the fiscal externality, which captures the elasticity of the tariff base with respect to the tariff itself. That elasticity is higher if higher tariffs induce agents to switch suppliers, or if agents avoid or evade the tariff via avenues like transshipment. What steps can the government take to reduce avoidant transshipment and thereby reduce the welfare cost of tariffs?

Traditionally, the public finance literature recommends traditional remedies like higher penalties, closer inspections, and more stringent audits (Allingham and Sandmo 1972; Slemrod and Yitzhaki 2002; Bhandari et al. 2024; Slemrod 2019). I argue—and show theoretically—that making imported intermediate and investment goods tax deductible on the domestic side is isomorphic to these more traditional mechanisms. Deductibility lowers the importer's effective burden of the tariff, making buyers less price-sensitive; exporters can then pass through more of the tariff on compliant shipments. That raises the profitability of the legitimate route relative to transshipping and reduces diversion, an effect that is behaviorally equivalent to tightening enforcement. This theoretical insight has immediate practical relevance: the One Big Beautiful Bill Act does precisely this by making all imported goods fully tax-deductible for producers.

2 Baseline Model

I begin with a partial equilibrium model in which heterogeneous exporters choose whether to ship directly to a country imposing a tariff, or to avoid the tariff via transshipment. Given that choice, they decide on a "free on board" (FOB) price. The key friction is a fixed cost associated with avoidance. Exporters differ in marginal cost, and demand on the domestic market is governed by a general, potentially non-CES, structure. Our focus is on the pricing behavior of exporters and how tariff-induced selection interacts with heterogeneity in passthrough behavior.

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.

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

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

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: $\dfrac{\partial D_i}{\partial \tilde{p}_i} < 0$ *and* $\dfrac{\partial D_i}{\partial \mathbf{P}} > 0$.

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

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

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

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

(b) $\kappa'(\tilde{p}) \; \geq \; -\frac{\varepsilon(\tilde{p})}{\varepsilon(\tilde{p})-1}\,\kappa(\tilde{p})^2.$

D1 ensures that demand slopes down in own delivered price and that higher rival prices (the aggregator $\mathbb{P}$) 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, ($\partial \kappa / \partial \ln \tilde{p} \leq 0$), a shape restriction that ensures passthrough among survivors is (weakly) increasing in delivered price—and hence in marginal cost.

Firm problem. Each firm $i \in \mathcal{J}$ 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

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

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

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

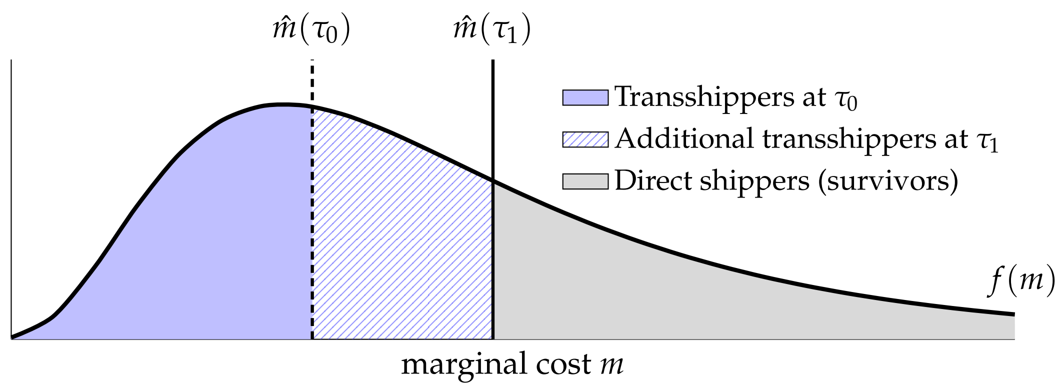
Moreover: (i) p_i^* strictly increases in m_i; (ii) in the direct channel D, p_i^* is weakly decreasing in τ (constant under CES); and (iii) q_i^* strictly decreases in m_i and, for D, in τ.
Proof: Appendix A.1

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 FOB 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 FOB 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, I first give some intuition for what follows. The story is effectively defined by Figure 1. Suppose, as I show formally later, that there exists a unique cutoff $\hat{m}(\tau)$. To the left of $\hat{m}(\tau)$, high-volume, low-cost firms choose to pay the fixed cost. To the right, firms choose to pass on part of the tariff to consumers. When the tariff rises, the cutoff moves to the right and expands the number of firms which choose to transship rather than try to pass on the fixed cost. Importantly, as long as markups are heterogeneous, direct shippers are selected for their ability to pass on the cost to domestic importers.

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.

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.

Formally, define the channel values

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

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

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 rises 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*

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

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

Proof. *Appendix A.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 among direct survivors.

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

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

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

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

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 our concavity primitive (**D4**), which is equivalent to $\varepsilon - 1 > \kappa$ at the optimum.[2] In the CES benchmark ($\kappa \equiv 0$), passthrough is homogeneous: $\beta_i = (1+\tau)^{-1}$ for all survivors.

The next result summarizes what matters for the rest of the paper—homogeneity under CES versus curvature-induced heterogeneity under Kimball-type demand.

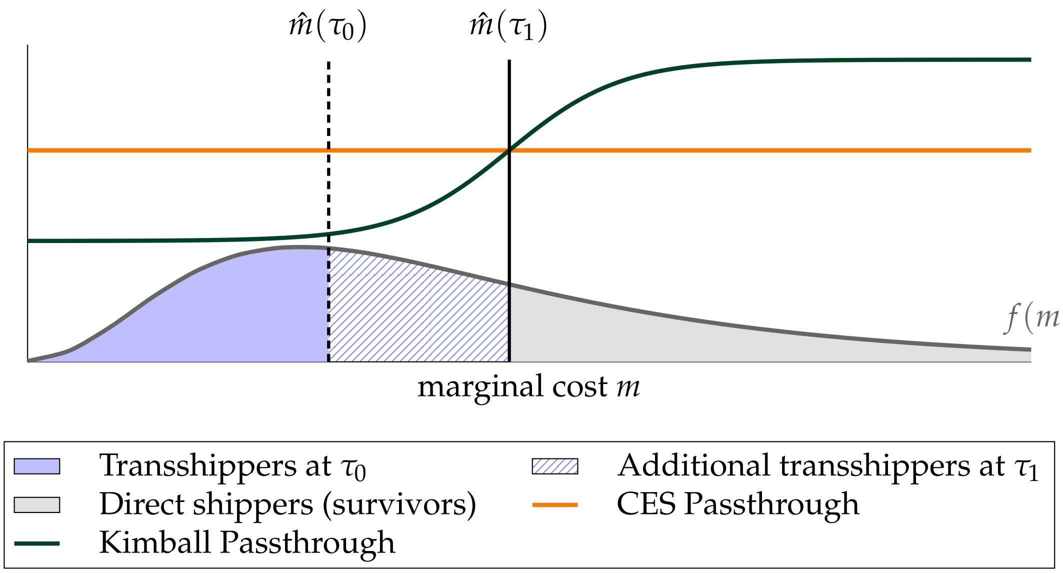
Proposition 2 (CES vs. Kimball among survivors)**.** *Fix P and assume D1–D4. (i)* CES: *If demand is CES ($\kappa \equiv 0$), then $\partial p_i^*/\partial\tau = 0$ and $\beta_i = (1+\tau)^{-1}$ for every survivor i. (ii) Kimball demand: If, in addition, D5a-b holds, then $\partial p_i^*/\partial\tau \leq 0$ 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 does too, so the average markup of survivors rises as well when markups are heterogeneous.

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 pass-through; 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.

2. A short derivation and the equivalence to **D4** are provided in Appendix A.

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 $\bar{\beta}$.

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 Selection and the Border Estimand

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

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

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:

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

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*

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

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

Two immediate corollaries follow. First, under CES ($\kappa \equiv 0$), all survivors have the same firm-level passthrough $\beta_i = (1+\tau)^{-1}$, so $\mathrm{Cov}(\beta_i, S_i) = 0$ 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 $\mathrm{Cov}(\beta_i, S_i) > 0$ 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—whereas 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.

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,

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

so any CPI or real-income calculation that is linear in passthrough should be based on β_{cohort}, not the border estimate. Using our selection identity and diversion share θ, a

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.

conservative estimate is given by

$$\beta_{\text{cohort}} \in \left[(1-\theta)\,\beta_{\text{border}},\ \beta_{\text{border}}\right],$$

implying that border-based incidence measures overstate the domestic price impact by up to a factor $1/(1-\theta)$ when $\theta > 0$.

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 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.

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).[4]

3.2 How much bias is there?

A simple metric for the magnitude of bias is the ratio of transshipped value to the pre-tariff value of targeted Chinese exports. Identifying transshipment is challenging; Freund et al. (2024) documents avoidance margins but does not quantify them, Iyoha et al. (2025) estimates flows through Vietnam alone, and Freund (2024) provides a multi-country estimate for 2022. The latter paper provides a methodology for identifying upper and lower

4. Cavallo et al. (2021) document a puzzle: exchange rate passthrough is far lower than measured tariff passthrough. It is likely that two factors cause this discrepancy. First, exchange rate shocks are often common across origins. Second, exchange rate shocks may easily reverse. Third, exchange rate shocks are transitory relative to permanent tariff policy. All three factors push against firms wanting to develop the infrastructure to engage in transshipment, which means that there is probably little bias in exchange rate passthrough regressions.

bounds to transshipment volumes. I use that as a guide for construct a time series estimate of total bias using annual data from the U.N. Comtrade database over the years 2017-2023.

We begin by restricting 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 diversion 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.

Both screens are implemented annually for 2018–2023 with 2017 as the baseline year. All flows are deflated to 2017 dollars using an import price index. Given that, I divide the upper and lower bounds by the 2017 value of exports from China which were targeted by tariffs. That yields a plausible range of estimates for the share of transshipped goods, denoted as $\left[\theta_{\text{low}}, \theta_{\text{high}}\right]$. With that, we can recover a corrected passthrough estimate by taking

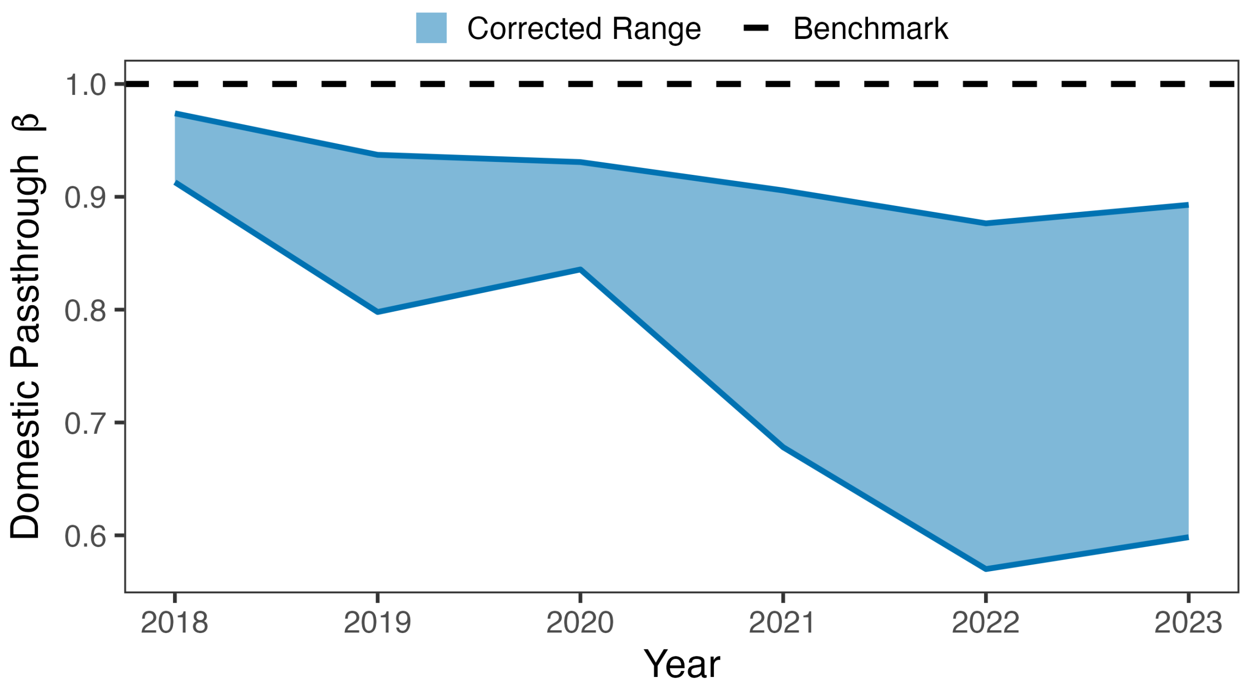
$$\beta_{\text{cohort}} \in \left[(1-\theta_{\text{high}})\beta_{\text{border}}, (1-\theta_{\text{low}})\beta_{\text{border}}\right].$$

I set $\beta_{\text{border}} = 1$ since border estimates tend to find complete passthrough. Figure 3 plots the resulting corrected passthrough estimate. In 2018, the estimated passthrough coefficient is between 90-95%, but it rapidly shrinks as time goes by. That makes sense because the tariffs were only in place for part of 2018 and it would take time for exporters

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.

to set up transshipment routes. Nevertheless, I view the true diversion share as closer to θ_{low}. The liberal screen likely overstates avoidance by classifying flows that reflect relocation, FDI in third-country production, or supplier switching as transshipment. These structural shifts leave the targeted tariff base but do not represent pure avoidance of the policy. Taken together, these considerations point to a true diversion share near θ_{low}, consistent with a working β_{cohort} around 0.85 when $\beta_{\text{border}} \approx 1$.

Figure 3: Corrected domestic passthrough bounds, 2018–2023



Note: Shaded band shows $[(1-\theta_{\text{high}}),\ (1-\theta_{\text{low}})]$ with θ from liberal and conservative screens applied to tariffed units; dashed line is the border benchmark $\beta_{\text{border}} = 1$. See the main text and Freund (2024) for construction details.

4 Enforcement via Tax Deductibility

Having shown that transshipment is empirically important, I now turn to welfare and enforcement.

4.1 Welfare implications

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 taxable income.[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 base. 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 0.7-1.4.[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 0.7-1.4 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 Expensing Intermediates as an Avoidance Reduction 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 diversion 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 valorem wedge $t_D \geq 0$ and the transshipment channel a (weakly) smaller wedge $t_T \in$

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.

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.

$[0, t_D)$. 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 $1+t$:

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

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* $\hat{m}(t_D, t_T) \in [\underline{m}, \bar{m}]$ *such that*

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

If the cutoff is interior, then

$$\frac{\partial \hat{m}}{\partial t_D} > 0 \qquad \textit{and} \qquad \frac{\partial \hat{m}}{\partial t_T} < 0.$$

Proof: See Appendix A.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 diversion mechanically.

Deductibility as raising the transshipment wedge. Suppose the buyer is a domestic firm facing corporate rate $\tau_c \in [0, 1)$ with full deductibility of tariff-inclusive input costs. After-tax unit costs are

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

so the direct–avoid cost gap under deductibility is

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

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

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

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*

$$t_T = \tau_c \, t_D,$$

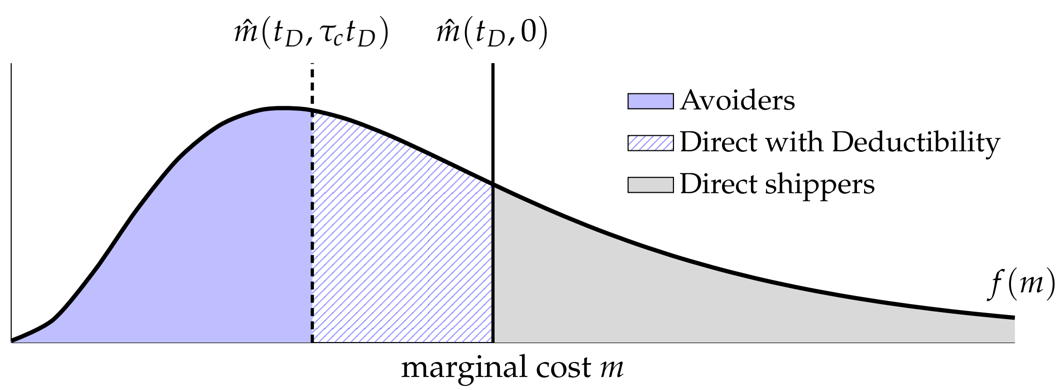
and with no deductibility the benchmark is $t_T = 0$.

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

$$\frac{d\hat{m}}{d\tau_c} = \frac{\partial \hat{m}}{\partial t_T} \cdot \frac{\partial t_T}{\partial \tau_c} = \underbrace{\frac{\partial \hat{m}}{\partial t_T}}_{<0} \cdot \underbrace{t_D}_{>0} < 0.$$

Higher deductibility ($\tau_c \uparrow$) 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 τ_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]

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.

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

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.

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, both components of the MCPF move favorably: domestic incidence uses $\gamma = \beta_{\text{cohort}}$, which falls with θ, and the fiscal externality falls as diversion shrinks. Deductibility therefore acts like an enforcement tool that reduces transshipment, aligns border measures with the cohort object, tilts revenue toward customs, and lowers the marginal welfare cost of tariff revenue.

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

Transshipment has first-order implications for measurement and policy. Empirically, rerouting drops the firms most likely to absorb the tariff from the border sample. As a result, border-based price effects should be read as upper bounds. Even modest rerouting—about 15–20 percent—can push border estimates up by a similar amount. That reduces the welfare cost of tariffs, but increases the fiscal externality. Policy can minimize the externality by narrowing the gap between complying and avoiding. Letting companies deduct the full cost of imported inputs makes avoidance less attractive, reduces diversion, and brings border measures closer to the unbiased estimates. That matters for understanding the evolution of tariffs, passthrough, and tax policy since equipment expensing began to phase out in 2023. Since tariffs were not fully deductible between 2023 and passage of OBBBA in July 2025, passthrough estimates made between Liberation Day and OBBBA are likely considerably more biased than estimates following OBBBA.

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.

Albrecht, Brian C., Thomas M. Phelan, and Nick Pretnar. 2023. *Time Use and the Efficiency of Heterogeneous Markups.* Technical report. Federal Reserve Bank of Cleveland, November. https://doi.org/10.26509/frbc-wp-202328.

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.

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.

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 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.

Freund, Caroline, Aaditya Mattoo, Alen Mulabdic, and Michele Ruta. 2024. "Is US trade policy reshaping global supply chains?" *Journal of International Economics* 152 (November): 104011. ISSN: 00221996. https://doi.org/10.1016/j.jinteco.2024.104011.

Gopinath, Gita, and Oleg Itskhoki. 2010. "Frequency of Price Adjustment and Pass-Through ¡sup¿*¡/sup¿." *Quarterly Journal of Economics* 125, no. 2 (May): 675–727. ISSN: 0033-5533. https://doi.org/10.1162/qjec.2010.125.2.675.

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.

A Proofs

A.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 first-order condition. Write $u \equiv \tilde{p}_i^c$ and $s \equiv \partial u/\partial p_i \in \{1, 1+\tau\}$. The FOC is

$$0 = \frac{\partial \Phi_i}{\partial p_i} = D_i(u) + (p_i - m_i)\, D_i'(u) \cdot s.$$

Using $\varepsilon_i(u) \equiv -\, uD_i'(u)/D_i(u)$, we obtain

$$\frac{p_i^* - m_i}{p_i^*} = \frac{u}{sp_i^*} \cdot \frac{-D_i(u)}{uD_i'(u)} = \frac{1}{\varepsilon_i(u)}.$$

This is the Lerner condition.

Monotonicity in m_i. Define $g(p, m; s) \equiv D(sp) + (p - m)sD'(sp)$. At the optimum, $g(p^*, m; s) = 0$. Compute partials:

$$\frac{\partial g}{\partial m} = -\, sD'(sp) > 0 \quad (\mathbf{D1}), \qquad \frac{\partial g}{\partial p} = 2sD'(sp) + (p-m)s^2 D''(sp).$$

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

Monotonicity in τ. In the direct channel set $u = sp$ with $s = 1 + \tau$ and define

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

Then

$$\frac{\partial G}{\partial s} = \frac{\varepsilon'(sp)}{\varepsilon(sp)^2}\, p \ \geq 0, \qquad \frac{\partial G}{\partial p} = \frac{m}{p^2} + \frac{\varepsilon'(sp)}{\varepsilon(sp)^2}\, s.$$

Under Kimball-type demand, $\varepsilon'(\cdot) \geq 0$, so $\partial G/\partial s \geq 0$. Profit concavity (D4, equivalently $\varepsilon - 1 - \kappa > 0$ at the optimum) implies $\partial G/\partial p > 0$. By the implicit function theorem,

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

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

Quantities. Since $\tilde{p}^D = sp^*$ and $\partial(sp^*)/\partial s > 0$ while $D'(\cdot) < 0$, $q_i^* = D(\tilde{p}^D)$ strictly falls in s; similarly, q_i^* strictly falls in m because p^* rises in m and demand slopes down. This proves (i)–(iii). $\square$

A.2 Proof of Proposition 1

Proof. Define the value function in the direct channel

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

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 objective $(p-m)D((1+\tau)p)$ with choice p and parameters $(m, 1+\tau)$,

$$\frac{\partial^2}{\partial m\, \partial(1+\tau)} \Big[(p-m)D((1+\tau)p)\Big] = -\, p\, D'\big((1+\tau)p\big) \;>\; 0 \quad (\textbf{D1}),$$

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 $p^*(m,\tau)$,

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

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

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

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

A.3 Proof of Lemma 2

Lemma 2 (Passthrough under ad valorem tariffs)**.** *Let $s \equiv 1+\tau$ and $\tilde{p}_D \equiv s\, p^*(m;D)$. Define $\varepsilon(u) \equiv -uD'(u)/D(u)$ 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 first-order condition as $G(p,s) \equiv (p-m)/p - 1/\varepsilon(sp) = 0$. Totally differentiate and use

$$\frac{\partial G}{\partial s} = \frac{\varepsilon'(sp)}{\varepsilon(sp)^2}\, p, \qquad \frac{\partial G}{\partial p} = \frac{m}{p^2} + \frac{\varepsilon'(sp)}{\varepsilon(sp)^2}\, s.$$

Using $\kappa(u) = -u\,\varepsilon'(u)/\varepsilon(u)$, we obtain $\partial G/\partial s = -\kappa(\tilde{p}_D)/[s\,\varepsilon(\tilde{p}_D)]$ and $\partial G/\partial p = m/p^2 - \kappa(\tilde{p}_D)/[p\,\varepsilon(\tilde{p}_D)]$. Hence

$$\frac{dp^*}{ds} = -\frac{\partial G/\partial s}{\partial G/\partial p} = \frac{\kappa(\tilde{p}_D)}{s}\cdot\frac{p^*}{\varepsilon(\tilde{p}_D)-1-\kappa(\tilde{p}_D)},$$

where the last step uses the Lerner condition $m = p^*(\varepsilon-1)/\varepsilon$. Finally,

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

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

A.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:

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

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

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

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

$$\frac{\partial \Delta}{\partial t_D} = -\frac{\partial V(m, 1+t_D)}{\partial t_D} \ \ge\ 0, \qquad \frac{\partial \Delta}{\partial t_T} = \frac{\partial V(m, 1+t_T)}{\partial t_T} \ \le\ 0.$$

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