Two-sided Heterogeneity: New Implications for Input Trade*

Tomohiro Ara†

Fukushima University

December 12, 2022

Abstract

This paper develops a heterogeneous firm model to analyze selection effects at different production stages on trade-induced intra-industry resource reallocations. Using a two-country symmetric setting in which both inputs and final goods are costly to trade subject to selection, we show that the trade elasticity of intermediate goods is endogenously greater than that of final goods due to an extra adjustment in the extensive margin. We also show that the welfare gains from input trade liberalization are greater than those from output trade liberalization if and only if the domestic input share is smaller than the domestic output share.

Keywords: Two-sided heterogeneity, input trade, selection, vertical linkages
JEL Classification Numbers: F12, F14

*This study was conducted as a part of the Project "Analyses of Offshoring" undertaken at the Research Institute of Economy, Trade and Industry (RIETI). Part of this paper is extended and is published as Ara and Zhang (2020) but the work presented here is a much more generalized development and has not been published elsewhere. I would like to thank Joaquin Blaum for invaluable comments and discussions which have significantly improved the quality of the paper. The current version of the manuscript also greatly benefits from two anonymous referee reports as well as detailed advice from the editor, Davin Chor. Financial support from the Japan Society of the Promotion of Science under grant numbers 19K01599, 20H01492 and 20H01498 is gratefully acknowledged. Data sharing is not applicable to this paper as no new data were created or analyzed in this study.

†Faculty of Economics and Business Administration, Fukushima University, Fukushima 960-1296, Japan. *Email address*: tomohiro.ara@gmail.com

1 Introduction

How different is the impact of trade barriers on trade flows between intermediate goods and final goods? How large are the welfare gains from input trade liberalization relative to output trade liberalization? Although the literature has devoted enormous effort to developing new trade models, there has been little theoretical work contrasting the implications of input and output trade liberalization in settings where every production stage features selection among heterogeneous producers. This paper tries to fill this important gap in the literature by deriving a gravity equation of intermediate goods and relating the trade elasticity obtained from that gravity equation to the welfare gains from input trade liberalization.

There is mounting evidence suggesting that input trade has been growing faster than final goods trade and its share is increasingly greater in the world trade volume, due to "outsourcing" or "offshoring" that fragments production processes across the globe (Hummels et al., 2001; Hanson et al., 2005; Johnson and Noguera, 2012). Recent empirical work has revealed that firms that import intermediate goods exhibit a number of the same performance differences as firms that export final goods. A series of work by Bernard et al. (2007, 2012, 2018a) unveils empirical regularity that, just as with exporting firms, importing firms are larger and more productive than non-importing firms in the same industry and only a small fraction of productive firms import inputs. Despite similarity in firm exporting and importing, another line of empirical work has shown that input trade liberalization has a quite different impact from output trade liberalization. Amiti and Konings (2007) find that input tariff reductions increase industry productivity more than twice as much as output tariff reductions in Indonesia; similarly, Topalova and Khandelwal (2011) find that firms' gains from input tariff reductions can be ten times greater than those from output tariff reductions in India. Moreover, input tariff reductions generate different productivity gains from output tariff reductions by expanding the technological possibilities of firms, as documented by Goldberg et al. (2010).

This paper develops a heterogeneous firm model to analyze selection effects at different production stages on trade-induced intra-industry resource reallocations. There is an industry consisting of two production stages (i.e., upstream and downstream stages) in our model where the former produces and exports intermediate goods by using labor, and the latter produces and exports final goods by using domestic/foreign inputs as well as labor. While suppliers at the upstream stage are modeled in a similar way to Melitz (2003), one of the main departures is that firms at the downstream stage must incur additional fixed costs when using foreign inputs in production, which may differ between markets to which they provide final goods. As a result, selection occurs not only at the upstream stage but also at the downstream stage so that only a fraction of productive firms use foreign inputs to serve the domestic and foreign markets. Our model featured with such two-sided heterogeneity then identifies potential channels through which trade liberalization in different types of goods has different impacts on trade-induced resource reallocations. For analytical simplicity, we study a symmetric-country setting; nonetheless, the model is able to highlight a new mechanism by focusing on selection of firms and suppliers at different stages, which jointly yields theoretical predictions in a consistent manner with empirical evidence.

Our first contribution is to show that an elasticity of the trade value with respect to trade barriers (referred to as the trade elasticity hereafter) is greater for intermediate goods than for final goods. The trade elasticity results are clearly seen in a special case of a Pareto distribution with free entry at the two production stages. Let T_j^{AB} denote the total value of trade flows between country A and country B where j is an index of the type of goods, which takes one of two forms: intermediate goods (M) or final goods (X). Then T_j^{AB} is given by

$$T_j^{AB} = \psi_j \times \left(Y^A\right)^a \times \left(Y^B\right)^b \times \left(\tau_j^{AB}\right)^{-\varepsilon_j}, \tag{1}$$

where ψ_j is a constant term, Y^A, Y^B are each country's GDP and τ_j^{AB} denotes variable trade costs of goods j. As with a usual gravity equation, (1) shows that the value of trade is positively affected by the size of exporting and importing countries but is negatively affected by trade barriers between these countries. In that sense, the gravity equation applies to intermediate goods as well as final goods. We find, however, that the trade elasticity estimated from that equation is *endogenously* greater for intermediate goods than for final goods, i.e., $\varepsilon_M > \varepsilon_X$. The trade elasticity difference arises because input and output trade costs affect selection of firms and suppliers at the respective production stages to different degrees.

The reason is very simple. On the one hand, reductions in input trade costs allow not only suppliers at the upstream stage to export intermediate goods more easily, but also firms at the downstream stage to import these inputs used in final goods production more easily. As new and less productive suppliers (firms) start exporting (importing) at the upstream (downstream) stage, the effects of input trade liberalization on input trade flows are amplified relative to that in single-stage production settings. On the other hand, reductions in output trade costs induce only new firms at the downstream stage to start exporting because suppliers at the upstream stage do not import final goods for input production. This suggests that there is an extra adjustment in the set of importers (i.e., extensive margin) in input trade that is absent in final goods trade, which makes the input trade elasticity greater.[1] The finding could help us to better understand why input trade has been growing faster than final goods trade in globalization (Hummels et al., 2001; Hanson et al., 2005; Johnson and Noguera, 2012). To demonstrate our contribution more sharply, we consider bilateral trade flows between symmetric countries; however, the trade elasticity results would hold for trade flows between asymmetric countries by relying on an exogenously fixed wage between countries.

Our second contribution is to show that the welfare gains from input trade liberalization are greater than those from output trade liberalization if and only if the domestic input share is smaller than the domestic output share. The welfare results are again clearly seen in a special case of a Pareto distribution with free entry at the two production stages. Let λ_j denote the domestic share where j is an index of the type of goods as in (1). Then the changes in welfare $\widehat{W} \equiv dW/W$ associated with variable trade costs τ_j are given by

$$\widehat{W} = \widehat{\lambda}_j^{-\frac{1}{\varepsilon_j}}, \tag{2}$$

where $\widehat{\lambda}_j \equiv d\lambda_j/\lambda_j$ denotes the changes in the domestic share in goods j. Hence, the welfare changes associated with trade liberalization in each type of goods can be computed only from the the domestic share and the trade elasticity of each type of goods estimated from (1). The welfare results follow from observing that (2) implies $d\ln W = -(1-\lambda_j)d\ln \tau_j$, as will be shown in our analysis.

When heterogeneity is present, trade liberalization triggers resource reallocations, forcing the least productive firms (suppliers) to exit the downstream (upstream) stage, which directly relates to the welfare gains from trade. Since such two-sided reallocations take place on a different scale between input and output trade liberalization, it is generally difficult to figure out economic factors capturing the welfare changes associated with these different trade shocks. Yet, in the special case above, the welfare changes can be computed only from the two statistics, λ_j and ε_j, for each type of goods j, which means that the welfare gains from input and output trade liberalization are the same conditional on these statistics. In other words, the Arkolakis et al. (2012) welfare formula applies

[1] In the model with roundabout production, it is known that the presence of importing leads to an expansion of both exporters and importers (Gopinath and Neiman, 2014; Blaum, 2019). Although this channel is operative in our model with vertical production in the sense that reductions in output trade costs indirectly induce suppliers at the upstream stage to start exporting (through an expansion of exporting firms at the downstream stage), the direct selection effects at each production stage are so strong that the input trade elasticity is always greater than the output trade elasticity in equilibrium.

even to the model of two-sided heterogeneity in which intermediate goods are produced at the upstream stage.[2] While the results may not be surprising, note that not only is the trade elasticity but also the domestic share is far away from equality between the two types of goods in the real world: the share of input trade is greater than that of final goods, and thus the domestic input share is smaller than the domestic output share. Given this fact, the welfare evaluation holding the two kinds of domestic share equal might lead to inaccurate understanding of globalization where fragmentation of production processes plays a prominent role in improving welfare in each country. The finding that the welfare gains are greater for input trade liberalization under rapidly rising input trade is also consistent with empirical evidence that input tariff reductions increase industry productivity more than output tariff reductions, because such productivity improvements are typically associated with the welfare gains from input trade liberalization more than from output trade liberalization.

In the remainder of this section, we review the related literature that explores the effects of tradable inputs on resource reallocations, trade flows and welfare gains from trade, paying attention to the relationships with the trade elasticity results in (1) and the welfare results in (2).

Antràs et al. (2017) develop a multi-country sourcing model in which firms' productivity has an endogenous impact on the number of countries from which firms source inputs. They find that, when fixed costs of foreign sourcing are large enough to generate firm selection into importing at the downstream stage, the aggregate trade elasticity tends to be greater than the firm-level trade elasticity in the gravity equation of intermediate goods. The result is derived, however, by assuming that final goods are non-tradable without analytical solutions for the trade elasticity. Moreover, they use the Eaton-Kortum (2002) framework for sourcing intermediate goods, implying that inputs are produced under conditions of perfect competition and supplier selection into exporting is not operative at the upstream stage. In contrast, we explicitly consider vertical linkages between input trade and final goods trade where both types of goods are costly to trade across borders subject to selection.

Antràs and de Gortari (2020) build a multi-stage production model of global value chains (GVCs) in which the specialization pattern of participant countries within GVCs is endogenously determined with trade barriers. In particular, they show that relatively more central countries tend to have comparative advantage and specialize in relatively downstream stages, since trade costs accumulate (or "compound") along GVCs. This compounding effect gives rise to the trade elasticity results with a very different flavor: trade at more downstream stages is more sensitive to changes in trade costs.[3] The contrasting results are clarified by the following two explanations. First is the market structure: all producers are perfectly competitive at each stage in their model, while all producers are imperfectly competitive at each stage in our model. Together with heterogeneity at each stage, the love-of-variety effect from an input expansion can be sufficiently strong in our model so as to outweigh the compounding effect of trade costs. Second is the specialization pattern: all countries completely specialize in production at each stage in their model, while all countries incompletely specialize in production at each stage in our model. Such incomplete specialization helps weaken the compounding effect of trade costs in our model, as some intermediate goods are sourced from the domestic market.

Bernard et al. (2018b) develop a multi-country model with two-sided heterogeneity where not only do firms at the downstream stage but also suppliers at the upstream stage produce with different production efficiencies.

[2] There are two classes of models of input-output linkages. First is "roundabout" production where output is sold to consumers as final goods and to firms as intermediate goods. This is often used in the literature, including Arkolakis et al. (2012, Section IV). Second is "vertical" or "sequential" production where intermediate (final) goods are produced at the upstream (downstream) stage. While this is also found in the literature, most previous work builds on perfectly competitive models and any firm-level variables do not play a key role in trade-induced intra-industry resource reallocations. We will elaborate on this point in the literature review.

[3] In contrast, Johnson and Moxnes (2019) find results similar to ours in the GVC model: input trade is more sensitive to changes in trade costs than final goods trade. Their results are, however, fundamentally driven by the endogenous reorganization of GVCs, whereas our results are driven by the endogenous response of the extensive margin at different stages.

Assuming that final goods are non-tradable, they find the negative degree assortivity among firms and suppliers, i.e., more productive firms match with less productive suppliers and thus buy inputs from a larger set of suppliers. In this transaction, while suppliers self-select into exporting after bearing fixed costs, firm selection takes place through matching without bearing fixed costs. In practice, however, such fixed costs are empirically relevant for explaining why a small fraction of firms import (Kasahara and Lapham 2013; Halpern et al. 2015). Abstracting from matching that leads to the negative degree assortivity, we show that trade liberalization triggers two-sided reallocations as a result of self-selection among firms and suppliers at different stages, which turns out to be crucial to deriving the difference in trade elasticities estimated from the gravity equation in (1).

Regarding welfare gains, our welfare results bear some resemblance to those in Melitz and Redding (2014). They show that when non-traded final goods are produced from a sequence of traded intermediate goods, the welfare gains from trade are amplified through an increase in domestic productivity. Similar welfare results are also reported by Ossa (2015) and Antràs and de Gortari (2020) in settings of multi-sector production and GVCs respectively. All of these papers, however, analyze perfectly competitive markets at every production stage (or sector) and any firm-level variables do not play a key role in intra-industry resource reallocations that directly affect the welfare gains in the presence of heterogeneity. In a strictly sequential production setup, this paper focuses on two-sided heterogeneity and identifies potential channels through which trade liberalization can have a different welfare impact from previous work through selection at different stages of production.[4]

The welfare changes in (2) are the same as those first shown by Arkolakis et al. (2012); the main difference lies in how trade liberalization reduces trade costs of inputs and final goods. To allow for tradable intermediate goods, they employ roundabout production in which firms use output of other firms as inputs in production and trade costs are the same between intermediate goods and final goods (as these goods are interchangeable). Then, their results can be interpreted as the welfare gains associated with trade liberalization where trade costs decrease *proportionately* between these two types of goods. In this scenario, the effects of trade liberalization on the welfare gains are amplified through an input-output loop, captured by new parameters absent in (2), and many papers have quantitatively demonstrated that the welfare gains are indeed greater with input trade than with final goods only in various settings (e.g., Costinot and Rodríguez-Clare, 2014; Antràs and Chor, 2022). In contrast, we allow trade costs to differ between intermediate goods and final goods that are produced at different stages. Then, our results can be interpreted as the welfare gains associated with trade liberalization in which trade costs of *only* one type of goods decrease, holding trade costs of another type of goods fixed. In our scenario, the effects of trade liberalization on the welfare gains are captured by (2) without introducing new parameters even with tradable inputs as in Melitz and Redding (2014), and our exercise shows that whether the welfare gains are greater for input trade liberalization than for output trade liberalization depends only on one of the sufficient statistics, i.e., domestic share.

This paper is most closely related to Ara and Zhang (2020) which extends our model setup to a multi-sector, asymmetric-country setup. While the trade elasticity results look similar, the scope of the papers is different. Ara and Zhang (2020) focus on country asymmetry in order to empirically investigate our theoretical prediction; however complexity of the model forces them to impose a restriction that either type of goods is only tradable. In contrast, we develop a more general model where both intermediate goods and final goods are tradable (albeit between two symmetric countries) and address the welfare implications that arise from their joint interaction. Drawing upon their work, we later provide empirical evidence supporting the trade elasticity results.

[4] In that respect, our model with imperfect competition at every stage borrows some insights originating from the IO literature; see for example Ishikawa and Spencer (1999), Ghosh and Morita (2007) and Ara and Ghosh (2016). These papers find that two-stage production models can have different welfare/policy implications from those obtained in single-stage production models.

2 Setup

Consider a model in which two symmetric countries trade both inputs and final goods. The economy has one industry consisting of upstream and downstream stages. Suppliers at the upstream stage produce differentiated intermediate goods, whereas firms at the downstream stage produce differentiated final goods by using inputs. Both stages are characterized by monopolistic competition with free entry. Labor is the only factor of production and each country is endowed with L units of labor which is chosen as the numéraire. Throughout the analysis, we provide key equations of the model relegating detailed derivations to the online Appendix (Section B.1).

2.1 Consumers

Consumer preferences are represented by a CES utility function with an elasticity of $\sigma > 1$:

$$U = \left(\int_i q_i^{\frac{\sigma-1}{\sigma}} di \right)^{\frac{\sigma}{\sigma-1}},$$

where q_i is firm i's output quantity. All inputs are produced by suppliers at the upstream stage in our model without roundabout production, which implies that firm output is sold to consumers only as final goods but is not sold to other firms as inputs. Thus firms' output quantity coincides with consumers' final goods demand. Utility maximization yields consumers' final goods demand for variety i:

$$q_i = p_i^{-\sigma} P^{\sigma-1} R,$$

where p_i is firm i's output price, P is the CES price index, and R is consumers' aggregate expenditure. Defining an aggregate output $Q \equiv U$, we have $PQ = R$.

2.2 Firms

Firm technology is represented by linear combination of firms' productivity and the input bundle: $q_i = \varphi x_i$. The productivity level φ is randomly drawn from a fixed distribution $G(\varphi)$ with unbounded upper support, while the input bundle x_i is produced by a CES production function with an elasticity of $\sigma > 1$:

$$x_i = \left(\int_v x_{Div}^{\frac{\sigma-1}{\sigma}} dv + \mathbb{1}_{Mi} \int_v x_{Miv}^{\frac{\sigma-1}{\sigma}} dv \right)^{\frac{\sigma}{\sigma-1}},$$

where x_{Div} and x_{Miv} are firm i's domestic and foreign input quantities from supplier v,[5] and $\mathbb{1}_{Mi}$ is an indicator function which takes the value of one if firm i uses foreign inputs and zero otherwise. To facilitate the analysis, we follow Bernard et al. (2018b) in assuming that inputs are combined without hiring labor and the elasticity of substitution between inputs in firm technology is identical with the elasticity of substitution between final goods in consumer preferences; however the simplification would not affect the qualitative results of this paper. Cost minimization yields firm i's input demand for variety v:

$$x_{Div} = p_{Dv}^{-\sigma} c_i^{\sigma-1} e_i,$$
$$x_{Miv} = p_{Mv}^{-\sigma} c_i^{\sigma-1} e_i,$$

[5] The subscripts i and v are attached to relevant variables to firms and suppliers respectively. Then the firm's productivity of variety i should be denoted by φ_i, but we drop the variety index i from productivity to simplify the notation. For the same reason, the supplier's productivity of variety v (introduced in Section 2.3) is simply denoted by ϕ rather than ϕ_v.

where p_{Dv} and p_{Mv} are domestic and foreign input prices set by supplier v (common to all firms i) and

$$c_i = \left(\int_v p_{Dv}^{1-\sigma} dv + \mathbb{1}_{Mi} \int_v p_{Mv}^{1-\sigma} dv \right)^{\frac{1}{1-\sigma}},$$
$$e_i = \int_v e_{Div} dv + \mathbb{1}_{Mi} \int_v e_{Miv} dv,$$

where $e_{Div} = p_{Dv} x_{Div}$ and $e_{Miv} = p_{Mv} x_{Miv}$ are firm i's domestic and foreign input expenditures to supplier v. A few points are in order for these specifications. First, substituting x_{Div} and x_{Miv} into the CES production function and rearranging, firm i's total input expenditure is expressed as

$$e_i = \frac{c_i}{\varphi} q_i.$$

Thus, firm i's input expenditure increases with its output quantity q_i but decreases with its productivity level φ. Second, from the input pricing rule set by suppliers and selection into exporting among suppliers, it follows that the price index associated with the input bundle (referred to as firm i's unit costs hereafter) is given by

$$c_i^{1-\sigma} = c_D^{1-\sigma} (1 + \mathbb{1}_{Mi} \tau_M^{1-\sigma} \Delta), \tag{3}$$

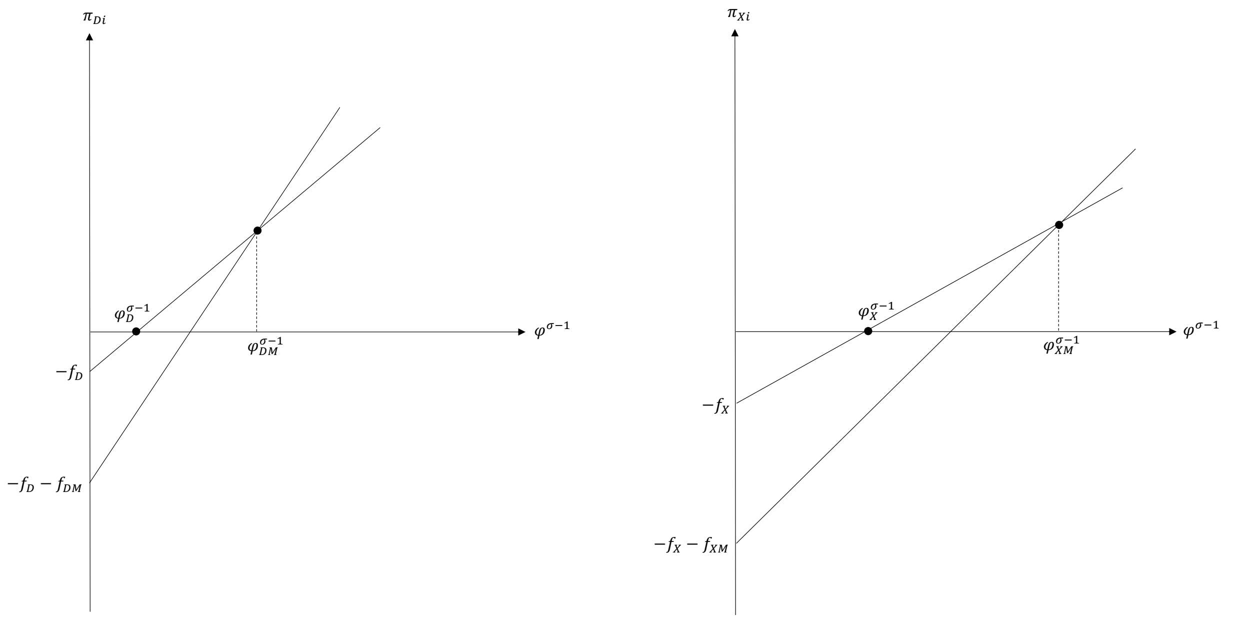
where $c_D^{1-\sigma} = \int_v p_{Dv}^{1-\sigma} dv$, τ_M denotes variable trade costs of inputs, and Δ is market share of exporting suppliers defined later. To understand (3), suppose that τ_M is sufficiently high that no supplier exports. Then $\Delta = 0$ and the unit costs are the same across firms. Evidence suggests however that firms using both domestic and foreign inputs have a cost advantage over firms using only domestic inputs (Halpern et al., 2015). Further, even if τ_M is not prohibitively high, there is selection into exporting at the upstream stage and not all suppliers export. Then $\Delta < 1$ and the unit costs of importing firms are lower than that of non-importing firms. Intuitively, firms using both kinds of inputs can exploit a love-of-variety effect in production and improve their production efficiency. Finally, the above specifications imply that when foreign inputs are sold to firms, all suppliers that export sell to all firms that import.[6] This can be seen from noting that (3) is independent of firms' productivity and all firms have the same set of suppliers: the difference in firms' unit costs comes solely from whether they import or not. This stands in contrast to recent evidence that more productive firms buy inputs from a larger set of suppliers (Bernard et al., 2018b) and probably (3) would not get support from the data. Recognizing the shortcomings, we assume this international matching between firms and suppliers, as it gives us a lot of tractability.

Given the sourcing strategy and associated unit costs, firm i's domestic profits are given by

$$\pi_{Di} = \left(p_{Di} - \frac{c_i}{\varphi} \right) p_{Di}^{-\sigma} P^{\sigma-1} R - f_{Di},$$

where f_{Di} denotes firm i's fixed overhead costs that satisfy $f_{Di} = f_D + \mathbb{1}_{Mi} f_{DM}$ (measured in units of labor). If firm i sources inputs only domestically, it incurs the fixed costs of domestic sourcing: $f_{Di} = f_D$. If firm i also sources inputs from abroad, in contrast, it incurs the *additional* fixed costs of foreign sourcing: $f_{Di} = f_D + f_{DM}$. This structure of fixed costs follows from the firm importing literature in which a firm who sources from abroad has to incur higher fixed costs of search, monitoring and communication (Antràs and Helpman, 2004). Firm i chooses its domestic price p_{Di} to maximize domestic profits π_{Di}. Noting that firm i takes the term $P^{\sigma-1}R$ as

[6] Another interpretation is that when a firm sources from abroad, it purchases a foreign bundle of inputs – so there is per se no matching – and this foreign bundle is a CES aggregate over all foreign input varieties currently available.



(a) Domestic market

(b) Foreign market

Figure 1 – Profits at the downstream stage

given, profit maximization implies that the firm sets output prices with a constant markup over marginal costs. This pricing rule in turn yields domestic revenues $r_{Di} = \sigma B c_i^{1-\sigma} \varphi^{\sigma-1}$ where

$$B = \frac{(\sigma-1)^{\sigma-1}}{\sigma^\sigma} P^{\sigma-1} R$$

is the index of final goods market demand. Using domestic revenues, domestic profits are then $\pi_{Di} = \frac{r_{Di}}{\sigma} - f_{Di}$. Comparing the unit costs c_i and the fixed costs f_{Di}, firm i chooses its domestic sourcing strategy so that

$$\pi_{Di} = \max\left\{0,\ Bc_D^{1-\sigma}\varphi^{\sigma-1} - f_D,\ B(1+\tau_M^{1-\sigma}\Delta)c_D^{1-\sigma}\varphi^{\sigma-1} - f_D - f_{DM}\right\}.$$

Figure 1(a) shows domestic profits. In the $(\varphi^{\sigma-1}, \pi_{Di})$ space, the slope of π_{Di} is $Bc_D^{1-\sigma}$ for non-importing firms and is $Bc_D^{1-\sigma}(1+\tau_M^{1-\sigma}\Delta)$ for importing firms, reflecting that variable profits are higher for importing firms due to lower unit costs. However, the intercept of π_{Di} is $-f_D$ for non-importing firms and is $-f_D - f_{DM}$ for importing firms, reflecting that fixed costs are higher for importing firms due to the additional sourcing costs. As a result, there are two productivity cutoffs at which domestic profits of non-importing firms are zero $\varphi_D^{\sigma-1}$, and domestic profits of importing firms exceed those of non-importing firms $\varphi_{DM}^{\sigma-1}$. This means that a fraction of more productive firms above $\varphi_{DM}^{\sigma-1}$ use both domestic and foreign inputs, whereas less productive firms between $\varphi_D^{\sigma-1}$ and $\varphi_{DM}^{\sigma-1}$ use only domestic inputs to serve the domestic market.

Among operating firms in the domestic market, some firms export final goods to serve the foreign market. The *additional* profits from exporting are given by

$$\pi_{Xi} = \left(p_{Xi} - \frac{\tau_X c_i}{\varphi}\right) p_{Xi}^{-\sigma} P^{\sigma-1} R - f_{Xi},$$

where f_{Xi} denotes firm i's fixed trade costs that satisfy $f_{Xi} = f_X + \mathbb{1}_{Mi} f_{XM}$ (measured in units of labor). In what follows, we allow the fixed costs to differ not only for non-importing firms ($f_D \neq f_X$) but also for importing firms ($f_{DM} \neq f_{XM}$). The latter implies that the additional fixed costs incurred when sourcing from abroad are different between domestic and foreign markets. While f_{DM} includes setup costs to build relations with foreign suppliers, f_{XM} includes such costs (which are included in f_{DM}) as well as coordination costs between exporting and importing activities (which are not included in f_{DM}). Since the tension between these considerations works contrariwise for the additional fixed costs, their ordering is generally ambiguous. For example, if economies of scope in diverse activities outweigh managerial overloads with more activities, f_{XM} can be smaller than f_{DM}.[7] We shall discuss the configuration of these fixed costs in detail in equilibrium analysis. In addition to the fixed trade costs, exporting firms incur variable trade costs of final goods denoted by τ_X and thus the unit costs are higher for exporting firms than for domestic firms.[8]

Profit maximization implies that firm i who exports sets higher prices due to increased marginal costs τ_X: $p_{Xi} = \tau_X p_{Di}$. This in turn yields export revenues $r_{Xi} = \sigma B (\tau_X c_i)^{1-\sigma} \varphi^{\sigma-1}$ and export profits $\pi_{Xi} = \frac{r_{Xi}}{\sigma} - f_{Xi}$. Comparing the unit costs $\tau_X c_i$ and the fixed costs f_{Xi}, firm i chooses its export sourcing strategy so that

$$\pi_{Xi} = \max \left\{ 0,\ B(\tau_X c_D)^{1-\sigma} \varphi^{\sigma-1} - f_X,\ B(1 + \tau_M^{1-\sigma} \Delta)(\tau_X c_D)^{1-\sigma} \varphi^{\sigma-1} - f_X - f_{XM} \right\}.$$

Figure 1(b) shows export profits, indicating a similar pattern between the domestic and foreign markets: there are two productivity cutoffs at which export profits are zero for non-importing firms $\varphi_X^{\sigma-1}$, and export profits of importing firms exceed those of non-importing firms $\varphi_{XM}^{\sigma-1}$. While a fraction of exporting firms above $\varphi_{XM}^{\sigma-1}$ use both domestic and foreign inputs as before, note that this partitioning is possible only if they incur the additional fixed costs f_{XM} to serve the foreign market.

To characterize equilibrium at the downstream stage, we need to identify the productivity cutoffs for firms. We denote them by φ_c where c is an index of the productivity cutoff for firms and thus $c \in \{D, X, DM, XM\}$. From Figures 1(a) and 1(b) and the domestic and export profits above, φ_c is given by

$$\begin{aligned} B c_D^{1-\sigma} \varphi_D^{\sigma-1} &= f_D, \\ B(\tau_X c_D)^{1-\sigma} \varphi_X^{\sigma-1} &= f_X, \\ B(\tau_M c_D)^{1-\sigma} \Delta \varphi_{DM}^{\sigma-1} &= f_{DM}, \\ B(\tau_X \tau_M c_D)^{1-\sigma} \Delta \varphi_{XM}^{\sigma-1} &= f_{XM}. \end{aligned} \tag{4}$$

Then, (4) shows the following selection patterns that arise at the downstream stage. First,

$$\left(\frac{\varphi_X}{\varphi_D} \right)^{\sigma-1} = \frac{\tau_X^{\sigma-1} f_X}{f_D}, \qquad \left(\frac{\varphi_{XM}}{\varphi_{DM}} \right)^{\sigma-1} = \frac{\tau_X^{\sigma-1} f_{XM}}{f_{DM}}. \tag{5}$$

If the variable trade costs τ_X and the fixed trade costs f_X, f_{XM} are so large that $\varphi_X > \varphi_D$ and $\varphi_{XM} > \varphi_{DM}$, (5) means firm selection into exporting, which holds not only for firms using only domestic inputs but also for firms using both domestic and foreign inputs; thus not all importing firms export. Moreover,

$$\left(\frac{\varphi_{DM}}{\varphi_D} \right)^{\sigma-1} = \frac{1}{\Delta} \frac{\tau_M^{\sigma-1} f_{DM}}{f_D}, \qquad \left(\frac{\varphi_{XM}}{\varphi_X} \right)^{\sigma-1} = \frac{1}{\Delta} \frac{\tau_M^{\sigma-1} f_{XM}}{f_X}. \tag{6}$$

[7]That the ordering of fixed costs depends on firms' activities is similar to that examined in Antràs and Helpman (2004).
[8]The subscripts M and X are attached to relevant variables to input trade and final goods trade respectively.

If the variable trade costs τ_M and the fixed trade costs f_{DM}, f_{XM} are so large that $\varphi_{DM} > \varphi_D$ and $\varphi_{XM} > \varphi_X$, (6) means firm selection into importing, which holds not only for the domestic market in Figure 1(a) but also for the foreign market in Figure 1(b); thus not all exporting firms import. These selection patterns accord well with empirical evidence that firm selection is ubiquitous for both exporting and importing (Bernard et al., 2007, 2012, 2018a). From these selection patterns, the productivity cutoffs at the downstream stage satisfy

$$\varphi_D < \min\left\{\varphi_{DM}, \varphi_X\right\} < \varphi_{XM}.$$

The ordering of the productivity cutoffs means that, among operating firms, those with the lowest productivity between φ_D and $\min\{\varphi_{DM}, \varphi_X\}$ use only domestic inputs and sell final goods in only the domestic market, while those with the highest productivity above φ_{XM} use both domestic and foreign inputs and sell final goods in both domestic and foreign markets.[9]

In addition to the zero profit cutoff condition, we impose a free entry condition. Upon bearing the fixed costs of entry f_E (measured in units of labor), a mass of entrants M_E draw their productivity φ from a distribution $G(\varphi)$. Firm i then decides whether to enter the downstream stage by choosing markets from which to source inputs and to which to provide final goods, or to exit without producing. Obviously the former outcome occurs whenever φ is greater than the domestic productivity cutoff φ_D, and hence the free entry condition is defined as $\int_{\varphi_D}^{\infty} \pi_i dG(\varphi) = f_E$ where the left-hand side represents firms' expected profits equivalent to $\frac{1}{M_E}\int_i \pi_i di$. This condition determines the final goods demand level B so that there are no pure profits at the downstream stage. Using the productivity cutoffs in (4), this condition can be expressed as

$$f_D J(\varphi_D) + f_X J(\varphi_X) + f_{DM} J(\varphi_{DM}) + f_{XM} J(\varphi_{XM}) = f_E, \tag{7}$$

where $J(\varphi_c) = \int_{\varphi_c}^{\infty}\left[\left(\frac{\varphi}{\varphi_c}\right)^{\sigma-1} - 1\right] dG(\varphi)$ is a strictly decreasing function of φ_c. It is worth emphasizing that the zero profit cutoff condition (4) and the free entry condition (7) cannot characterize the downstream stage. As in (6), the import productivity cutoffs φ_{DM}, φ_{XM} are affected by the market share of exporting suppliers Δ which is endogenously determined at the upstream stage. This shows that any trade shocks that induce changes in Δ at the upstream stage have an impact on firm selection into importing at the downstream stage through the availability of foreign inputs used in final goods production by firms.

We conclude this section by deriving the domestic output share as well as the domestic input share. First, from the fact that the aggregate revenue of firms equals the aggregate expenditure of consumers, the domestic output share (from the viewpoint of consumers) is given by

$$\lambda_X = \frac{\int_i r_{Di} di}{\int_i r_i di} = \frac{1}{1 + \tau_X^{1-\sigma}\Lambda_X}, \tag{8}$$

where

$$\Lambda_X = \frac{V(\varphi_X) + \tau_M^{1-\sigma}\Delta V(\varphi_{XM})}{V(\varphi_D) + \tau_M^{1-\sigma}\Delta V(\varphi_{DM})}$$

and $V(\varphi_c) = \int_{\varphi_c}^{\infty} \varphi^{\sigma-1} dG(\varphi)$ is a strictly decreasing function of φ_c. It can be shown that the numerator and denominator of Λ_X are proportional to $\int_i r_{Xi} di$ and $\int_i r_{Di} di$ respectively. Following Melitz and Redding (2015), Λ_X is referred to as the market share of exporting firms, though we recognize that there are many definitions of

[9] When the impact of trade liberalization is examined in Section 3, we restrict the main analysis to small changes in trade costs that preserve the ordering of the productivity cutoffs at the downstream stage.

market share in the literature. From this, if the variable trade costs of inputs are sufficiently high ($\tau_M = \infty$) so that no supplier exports ($\Delta = 0$), the domestic output share (8) collapses to that in the plain Melitz model.[10] Second, from the fact that the aggregate revenue of suppliers equals the aggregate input expenditure of firms, the domestic input share (from the viewpoint of firms) is given by

$$\lambda_M = \frac{\int_i e_{Di} di}{\int_i e_i di} = \frac{1}{1 + \tau_M^{1-\sigma} \Delta \Lambda_M}, \tag{9}$$

where

$$\Lambda_M = \frac{V(\varphi_{DM}) + \tau_X^{1-\sigma} V(\varphi_{XM})}{V(\varphi_D) + \tau_X^{1-\sigma} V(\varphi_X)}.$$

It can be shown that the numerator and denominator of Λ_M are proportional to $\int_i e_{Mi} di$ and $\int_i e_{Di} di$ respectively, and Λ_M is referred to as the market share of importing firms. Not surprisingly, if the variable trade costs of inputs are sufficiently high ($\tau_M = \infty$) so that no supplier exports ($\Delta = 0$) and no firm imports ($\Lambda_M = 0$), the domestic input share (9) collapses to unity. To make the analysis interesting, we restrict the range of τ_M under which $\Delta < 1$ and $\Lambda_M < 1$.

2.3 Suppliers

Supplier technology is represented by a linear cost function of labor that involves fixed costs and marginal costs where the latter are inversely related to suppliers' productivity:

$$l_v^p = k + \frac{x_v}{\phi}.^{11}$$

All suppliers incur the same fixed costs k which vary with the markets to which suppliers provide their inputs. The productivity level ϕ is randomly drawn from a fixed distribution $G(\phi)$ with unbounded upper support, while $x_v = \int_i x_{iv} di$ is the total input quantity provided by supplier v where x_{iv} is firm i's input demand for variety v derived in the last section. This follows from our matching assumption that all suppliers that export sell to all firms that import. These features of input production imply that the upstream stage is also characterized by monopolistic competition in which suppliers self-select into the domestic and foreign markets, as with firms at the downstream stage. While we assume that the productivity distribution of suppliers $G(\phi)$ is identical with that of firms $G(\varphi)$ below, this is not critical to our analysis and can be relaxed as shown later.

Given suppliers' production technology and firms' input demand, supplier v's domestic profits are given by

$$\pi_{Dv} = \left(p_{Dv} - \frac{1}{\phi}\right) p_{Dv}^{-\sigma} \int_i c_i^{\sigma-1} e_i di - k_D,$$

where k_D denotes suppliers' fixed overhead costs that are common to operating suppliers in the domestic market. Supplier v chooses its domestic price p_{Dv} to maximize domestic profits π_{Dv}. Noting that supplier v takes the term $\int_i c_i^{\sigma-1} e_i di$ as given, profit maximization implies that the supplier sets input prices with a constant markup over marginal costs. This pricing rule in turn yields domestic revenues $r_{Dv} = \sigma A \phi^{\sigma-1}$ where

$$A = \frac{(\sigma-1)^{\sigma-1}}{\sigma^\sigma} c_D^{\sigma-1} \int_i e_{Di} di$$

[10] See, for example, eq (18) in Melitz and Redding (2015) which can be thought of as the limiting case of ours in which $\tau_M = \infty$.

[11] The superscript p is attached to stress production worker, whereas the superscript e will be attached to labor used for entry, as shown in Section 2.4. The same applies to labor at the downstream stage.

is the index of input market demand. Then, domestic profits are rewritten as $\pi_{Dv} = \frac{r_{Dv}}{\sigma} - k_D$, and supplier v chooses its domestic production strategy so that

$$\pi_{Dv} = \max\left\{0,\ A\phi^{\sigma-1} - k_D\right\}.$$

As in Figure 1(a), domestic profits can be shown in the $(\phi^{\sigma-1}, \pi_{Dv})$ space where the slope of π_{Dv} is A and the intercept is $-k_D$. While this pins down the domestic productivity cutoff for suppliers as in firms, there is a crucial difference between firms and suppliers. Recall that firms decide not only whether to serve the domestic market but also whether to import inputs from abroad in their domestic strategy. In contrast, suppliers decide only whether to serve the domestic market in their domestic strategy since they do not use final goods in input production in our model with vertical production (i.e., without roundabout production). As a result, there is a unique productivity cutoff at which domestic profits of operating suppliers are zero, namely $\phi_D^{\sigma-1}$. This means that a fraction of more productive suppliers above $\phi_D^{\sigma-1}$ provide their inputs to the domestic market, whereas less productive suppliers below $\phi_D^{\sigma-1}$ immediately exit without producing.

Among operating suppliers in the domestic market, some suppliers export intermediate goods to serve the foreign market. The *additional* profits from exporting are given by

$$\pi_{Mv} = \left(p_{Mv} - \frac{\tau_M}{\phi}\right) p_{Mv}^{-\sigma} \int_i c_i^{\sigma-1} e_i di - k_M,$$

where k_M denotes suppliers' fixed export costs that are common to exporting suppliers. In addition to the fixed trade costs, exporting suppliers incur variable trade costs of inputs denoted by τ_M. While trade costs that suppliers incur at the upstream stage are similar to those that firms incur at the downstream stage, there are no additional fixed costs of foreign sourcing for suppliers.

Profit maximization implies that supplier v who exports sets higher prices due to increased marginal costs τ_M ($p_{Mv} = \tau_M p_{Dv}$), just as firm i who exports sets higher prices due to increased marginal costs τ_X ($p_{Xi} = \tau_X p_{Di}$). From firms' unit costs in (3), this implies that the markup on the input prices (set by suppliers) is fully passed through to the output prices (set by firms) bundled from these inputs, i.e., a double marginalization takes place between firms and suppliers.[12] Combining the input pricing rule and the foreign input demand, we get export revenues $r_{Mv} = \sigma A \tau_M^{1-\sigma} \Lambda_M \phi^{\sigma-1}$. It is important to note that export revenues r_{Mv} include the market share of importing firms Λ_M. Not surprisingly, the input trade costs τ_M decrease export revenues by making it difficult to export their inputs. This direct effect of input trade costs is captured by $\tau_M^{1-\sigma}$ in export revenues. When there is firm selection into importing at the downstream stage, however, the input trade costs also decrease the mass of importing firms via the pass-through of the markup described above. This indirect effect is captured by Λ_M in export revenues. The effect of input trade costs on the two production stages can be clearly seen in the domestic input share (9), in the sense that the input trade costs affect the market share of importing firms Λ_M as well as the market share of exporting suppliers Δ. Since export profits are expressed as $\pi_{Mv} = \frac{r_{Mv}}{\sigma} - k_M$, supplier v chooses its export strategy so that

$$\pi_{Mv} = \max\left\{0,\ A\tau_M^{1-\sigma}\Lambda_M\phi^{\sigma-1} - k_M\right\}.$$

[12] This pass-through of the markup in the model with vertical production is very similar to that explored in Caliendo et al. (2021) in the model with roundabout production.

Like Figure 1(b), export profits can be shown in the $(\phi^{\sigma-1}, \pi_{Mv})$ space. It is straightforward to see that there is a unique productivity cutoff, namely $\phi_M^{\sigma-1}$, above which suppliers profitably export their inputs to the foreign market.

To characterize equilibrium at the upstream stage, we need to identify the productivity cutoffs for suppliers. We denote them by ϕ_c where, with some abuse of notation, c is an index of the productivity cutoff for suppliers as well and thus $c \in \{D, M\}$ in this case. From the domestic and export profits above, ϕ_c is given by

$$\begin{aligned} A\phi_D^{\sigma-1} &= k_D, \\ A\tau_M^{1-\sigma}\Lambda_M\phi_M^{\sigma-1} &= k_M. \end{aligned} \tag{10}$$

Then, (10) shows the following selection pattern at the upstream stage:

$$\left(\frac{\phi_M}{\phi_D}\right)^{\sigma-1} = \frac{1}{\Lambda_M}\frac{\tau_M^{\sigma-1}k_M}{k_D}. \tag{11}$$

If the variable trade costs τ_M and the fixed trade costs k_M are so large (while keeping $\Lambda_M < 1$) that $\phi_M > \phi_D$, (11) means supplier selection into exporting.[13] Note importantly similar selection patterns between (6) and (11). Whereas (6) imposes firm selection into importing at the downstream stage, (11) imposes supplier selection into exporting at the upstream stage. This is consistent with the previous explanation that the input trade costs simultaneously affect the mass of importing firms as well as the mass of exporting suppliers in our model with vertical production.

In addition to the zero profit cutoff condition, we impose a free entry condition. Upon bearing the fixed costs of entry k_E (measured in units of labor), a mass of entrants N_E draw their productivity ϕ from a distribution $G(\phi)$. Supplier v then decides whether to enter the upstream stage by choosing markets to which to provide intermediate goods or to exit without producing. The former outcome occurs whenever ϕ is greater than the domestic productivity cutoff ϕ_D, and the free entry condition is defined as $\int_{\phi_D}^{\infty} \pi_v dG(\phi) = k_E$ where the left-hand side represents suppliers' expected profits equivalent to $\frac{1}{N_E}\int_v \pi_v dv$. This condition determines the input demand level A so that there are no pure profits at the upstream stage. Using the productivity cutoffs in (10),

$$k_D J(\phi_D) + k_M J(\phi_M) = k_E, \tag{12}$$

where the functional form of $J(\phi_c)$ is the same as that of $J(\varphi_c)$, as the productivity distributions are identical between firms and suppliers. As with the downstream stage, equilibrium conditions in (10) and (12) cannot characterize the upstream stage, as the export productivity cutoff ϕ_M is affected by the market share of importing firms Λ_M that is endogenously determined at the downstream stage. Thus the impact of trade liberalization cannot be examined without taking into account vertical linkages between these two production stages.

We conclude this section by deriving the domestic input share. Noting that the aggregate revenue of suppliers equals the aggregate input expenditure of firms, the domestic input share (9) is alternatively defined as the ratio of the aggregate revenue of domestic inputs to the aggregate revenue of total inputs earned by suppliers. Using the domestic and export revenues above, the domestic input share is given by

$$\lambda_M = \frac{\int_v r_{Dv} dv}{\int_v r_v dv} = \frac{1}{1 + \tau_M^{1-\sigma}\Delta\Lambda_M},$$

[13] The ordering of the productivity cutoffs at the upstream stage is also preserved for any trade shocks in the paper.

where

$$\Delta = \frac{V(\phi_M)}{V(\phi_D)},$$

and the functional form of $V(\phi_c)$ is the same as that of $V(\varphi_c)$. We can show that the numerator and denominator of Δ are proportional to $\int_v r_{Mv} dv$ and $\int_v r_{Dv} dv$ respectively, and Δ is referred to as the market share of exporting suppliers. Observe that $\Delta < 1$ so long as there is supplier selection into exporting at the upstream stage in (11). The firm's unit costs in (3) are obtained from this market share and the input pricing rule set by suppliers.

2.4 Labor Market and Welfare

To close the model, we impose the labor market clearing condition. Noting that labor is used in both downstream and upstream stages, the condition is expressed as

$$\int_i l_i di + \int_v l_v dv = L,$$

where $l_i = l_i^e + l_i^p$ and $l_v = l_v^e + l_v^p$ are labor used for entry and production by firm i and supplier v respectively. Substituting labor used by firms and suppliers, the labor market clearing condition is simply expressed as $R = L$, and hence the aggregate revenue of firms R equals the aggregate payment to labor L in the economy. Moreover, the aggregate amount of labor used at each production stage is respectively given by

$$\int_i l_i di = \frac{L}{\sigma}, \quad \int_v l_v dv = \left(\frac{\sigma - 1}{\sigma}\right) L,$$

which shows that the labor allocation between the two production stages is exogenously fixed. Using the labor market clearing condition as well as the zero profit cutoff and free entry conditions, the mass of entrants at the two production stages M_E, N_E is written as a function of the labor endowment L and the productivity cutoffs φ_c, ϕ_c at the respective production stage.

Welfare per worker is defined as an inverse of the CES price index. In our vertical production model where the output prices are influenced by the input prices, it depends on all types of firms' productivity cutoffs (4) as well as all types of suppliers' productivity cutoffs (10). Using the labor market clearing condition, however, the dependence of the price index on the foreign productivity cutoffs is eliminated, and welfare depends only on the domestic productivity cutoffs at the two production stages. In fact, we can express welfare as

$$W = \sigma^{\frac{2}{1-\sigma}} \left(\frac{\sigma - 1}{\sigma}\right)^{\frac{2\sigma-1}{\sigma-1}} L^{\frac{2}{\sigma-1}} (f_D k_D)^{-\frac{1}{\sigma-1}} \varphi_D \phi_D \lambda_M^{\frac{1}{\sigma-1}}. \tag{13}$$

As usual in the literature, welfare is higher in a larger country due to increased product variety. (13) also shows that a larger population increases welfare with an elasticity of $2/(\sigma - 1)$ in our two-stage production model, while it increases welfare with an elasticity of $1/(\sigma - 1)$ in the single-stage production model with free entry, i.e., the presence of vertical linkages can amplify the effects of country size on welfare via an input expansion. More important are the sufficient statistics for welfare: when both inputs and final goods are costly to trade subject to selection, welfare is determined not only by firms' domestic productivity cutoff φ_D at the downstream stage, but also by suppliers' domestic productivity cutoff ϕ_D and the domestic input share λ_M at the upstream stage.

This completes the characterization of the model. We address the impact of trade on resource reallocations, trade flows, and welfare gains by solving for general equilibrium of the model.

3 Equilibrium

There are the two zero profit cutoff conditions (4), (10) and the two free entry conditions (7), (12) in our model. As we have seen in Sections 2.2 and 2.3, the former determine the productivity cutoffs at each production stage $(\phi_D, \phi_M, \varphi_D, \varphi_X, \varphi_{DM}, \varphi_{XM})$, while the latter determine the demand levels at each production stage (A, B). The conditions jointly provide implicit solutions for the above eight unknowns in general equilibrium where the labor market clearing condition is omitted by choosing labor as the numéraire. It is important to note that the assumption of country symmetry comes in, which enables us to obtain the sharp analytical results in the paper. Once these unknowns are determined, other endogenous variables of the model are automatically determined.

3.1 Resource Reallocations

We start by examining the impact of trade liberalization on the productivity cutoffs. Recall that the market share of exporting suppliers Δ and that of importing firms Λ_M enter the zero profit cutoff conditions in (4) and (10) respectively. Then, changes in the equilibrium variables associated with trade liberalization depend on how these market shares are affected by such liberalization. For example, totally differentiating $\Delta = V(\phi_M)/V(\phi_D)$, changes in the market share of exporting suppliers at the upstream stage are given by

$$d \ln \Delta = -\theta_M d \ln \phi_M + \theta_D d \ln \phi_D,$$

where $\theta_c \equiv -d \ln V(\phi_c)/d \ln \phi_c$ can be thought of as the extensive margin elasticity at the upstream stage.[14] By definition, the extensive margin elasticity θ_c is a function of the productivity cutoff ϕ_c; thus changes in Δ come not only from changes in ϕ_c but also from changes in $V(\phi_c)$, i.e., extensive margin elasticity θ_c. To make the analysis as simple as possible, we hereafter restrict attention to a subset of the general productivity distributions where the extensive margin elasticity satisfies $\theta_c = \theta$ for any c. This means that the extensive margin elasticity is the same across all productivity cutoffs taking a constant value regardless of suppliers' global status. In a similar vein, let $\vartheta_c \equiv -d \ln V(\varphi_c)/d \ln \varphi_c$ denote the extensive margin elasticity at the downstream stage and we assume that $\vartheta_c = \theta$ for any c. Admittedly, this is a restrictive assumption but it holds under one of the most commonly-used distributions in the literature: Pareto. More specifically, if ϕ and φ are distributed Pareto with a common shape parameter γ, the extensive margin elasticities at the two production stages are given by

$$\theta_c = \vartheta_c = \theta = \gamma - (\sigma - 1). \tag{14}$$

Although recent empirical work reports that the extensive margin elasticities are less likely to be constant (e.g., Bas et al., 2017), the Pareto distribution is nonetheless a good approximation of observed micro-level data and we follow standard practice in the literature.

Before proceeding further, we emphasize that the following analysis does *not* hinge on assuming a common Pareto shape parameter shared by upstream and downstream stages. Indeed the key analytical results can be generalized to allow for a different shape parameter between these two production stages, i.e., $\theta \neq \vartheta$. Relaxing this assumption, we can also study how the degree of heterogeneity across stages affects equilibrium outcomes. In this setting, however, calculations are cumbersome as two separate shape parameters need to be introduced into the analysis. To make our paper's point sharper, we focus on the simple setting for now. Any new results that arise in such a more general setting will be discussed in Section 3.4.

[14] See Arkolakis et al. (2012, p.110). In our model, θ_c is the extensive margin elasticity associated with ϕ_c for $c \in \{D, M\}$.

Input Trade Liberalization. We first consider the effect of input trade liberalization, i.e., reductions in input trade costs τ_M holding output trade costs τ_X fixed. Differentiating and solving the equilibrium conditions yields the following expressions of changes in the domestic productivity cutoffs ϕ_D, φ_D (see Appendix A.1 for proof):

$$d\ln\phi_D = d\ln\varphi_D = -\left(\frac{(\sigma-1)(1-\lambda_M)}{\sigma-1-\theta}\right) d\ln\tau_M. \tag{15}$$

(15) shows that changes in the domestic productivity cutoffs associated with input trade liberalization are the same between the two production stages, which strongly relies on assumption (14). Further, the least productive firms and suppliers are forced to exit the respective production stage by reductions in τ_M if and only if

$$\sigma - 1 > \theta. \tag{16}$$

In the inequality, $\sigma-1$ is the (common) intensive margin elasticity under the CES utility/production functions, while θ is the (common) extensive margin elasticity at the upstream/downstream stages under the distributional assumption. Hence, (16) requires that the extensive margin elasticity is not too large relative to the intensive margin elasticity at each production stage.[15] We impose this condition since input trade liberalization otherwise induces less productive firms and suppliers to enter the respective stage, which would be less likely in reality.

It is possible to examine changes in other productivity cutoffs. At the upstream stage, from the free entry condition (12), it follows that the export productivity cutoff ϕ_M shifts in the opposite directions to ϕ_D, and input trade liberalization generates the following changes in ϕ_D, ϕ_M:

$$\frac{d\ln\phi_D}{d\ln\tau_M} < 0 < \frac{d\ln\phi_M}{d\ln\tau_M}.$$

From these changes, we know that input trade liberalization leads to the Melitz-type resource reallocations at the upstream stage: more (less) productive suppliers gain (lose) their share in the economy. At the same time, input trade liberalization also generates changes in other productivity cutoffs at the downstream stage which operates through vertical linkages between these production stages. From the zero profit cutoff condition (5) and the free entry condition (7), changes in firms' productivity cutoffs associated with such liberalization are

$$\frac{d\ln\varphi_D}{d\ln\tau_M} = \frac{d\ln\varphi_X}{d\ln\tau_M} < 0 < \frac{d\ln\varphi_{DM}}{d\ln\tau_M} = \frac{d\ln\varphi_{XM}}{d\ln\tau_M}.$$

This means that resource reallocations arise even across exporting firms: more productive firms sourcing inputs from multiple markets expand by input trade liberalization, whereas less productive firms sourcing inputs from only a single market shrink by such liberalization. As a result, input trade liberalization leads exporting firms that import (do not import) inputs from abroad to gain (lose) their share in the economy.

Lemma 1: *If inputs and final goods are costly to trade across borders subject to selection under (16),*

(i) *Input trade liberalization gives rise to resource reallocations not only at the upstream stage but also at the downstream stage.*

(ii) *Input trade liberalization simultaneously induces such reallocations even across exporting firms: more (less) productive exporting firms sourcing inputs from multiple markets (a single market) expand (shrink).*

[15] The condition holds for parameter values used in the literature. For example, Melitz and Redding (2015) set $\sigma = 4$ and $\theta = 1.25$.

Output Trade Liberalization. Let us next consider the effect of output trade liberalization, i.e., reductions in output trade costs τ_X holding input trade costs τ_M fixed. This exercise allows us to explore the difference in resource reallocations between input and output trade liberalization. As shown in Appendix A.2, we have the following expressions of changes in the domestic productivity cutoffs ϕ_D, φ_D:

$$\begin{aligned} d\ln\phi_D &= -\left(\frac{(\sigma-1)(\mu_D-\mu_M)(1-\lambda_M)}{\sigma-1-\theta}\right) d\ln\tau_X, \\ d\ln\varphi_D &= -\left(1-\lambda_X+\frac{\theta(\mu_D-\mu_M)(1-\lambda_M)}{\sigma-1-\theta}\right) d\ln\tau_X, \end{aligned} \tag{17}$$

where μ_D (μ_M) is the domestic output share of firms using only domestic inputs (both domestic and foreign inputs).[16] Then (17) shows that, in contrast to (15), changes in the domestic productivity cutoffs associated with output trade liberalization are different between the two production stages even under assumption (14). The impact of output trade liberalization, however, is qualitatively similar to that of input trade liberalization, in the sense that the least productive firms and suppliers are forced to exit the respective stage by reductions in τ_X. This result requires both (16) and $\mu_D - \mu_M > 0$ where the latter inequality is equivalent to

$$\frac{V(\varphi_{XM})}{V(\varphi_{DM})} > \frac{V(\varphi_X)}{V(\varphi_D)}. \tag{18}$$

To understand (18), recall that $V(\varphi_c) = \int_{\varphi_c}^{\infty} \varphi^{\sigma-1} dG(\varphi)$ and each firm's revenues are proportional to $\varphi^{\sigma-1}$ in every market. Thus we can think of $V(\varphi_c)/V(\varphi_{c'})$ as summarizing the distribution of relative revenues of firms with productivity above any given cutoffs $\varphi_c, \varphi_{c'}$ in (4). In light of this, (18) implies that revenues of firms that simultaneously export and import are relatively larger than those of firms that export only or import only, i.e., complementarity between exporting and importing that happens via an increase in firm revenues (see Section B.2 in the online Appendix for detailed discussions). Moreover, since $V(\varphi_c)$ is a strictly *decreasing* function of φ_c where each productivity cutoff is proportional to relevant fixed costs, (18) may require that f_{XM} is sufficiently smaller than f_{DM}, and/or f_X is sufficiently larger than f_D. This intuition is actually correct when we posit the Pareto distribution that leads to (14). Noting $V(\varphi_c)/V(\varphi_{c'}) = (\varphi_{c'}/\varphi_c)^{\theta}$ under this specific parameterization, it is possible to express (18) in terms of fixed costs only:

$$\frac{f_{DM}}{f_{XM}} > \frac{f_D}{f_X}.$$

The configuration of fixed costs means that when a firm is already an importer, becoming an exporter does not entail full payment of fixed export costs, but rather a smaller amount of such fixed costs, i.e., complementarity between exporting and importing that happens via a reduction in fixed costs (which would not necessarily hold under a non-Pareto distribution). In fact, the theoretical result of (18) is consistent with recent empirical work documenting that intense exporters tend to be intense importers and their share is relatively larger than firms that export only or import only (Blaum, 2019). From these reasons, we will assume not only (16) but also (18) in the following analysis.

With these conditions being satisfied, we can see changes in other productivity cutoffs. On the one hand, from the free entry condition (12), ϕ_M shifts in the opposite directions to ϕ_D, giving rise to resource reallocations from less productive suppliers to more productive suppliers at the upstream stage, as with reductions in τ_M.

[16] $\mu_D = \frac{1}{1+\tau_X^{1-\sigma}\frac{V(\varphi_X)}{V(\varphi_D)}}$ and $\mu_M = \frac{1}{1+\tau_X^{1-\sigma}\frac{V(\varphi_{XM})}{V(\varphi_{DM})}}$. Note that these always satisfy $1 > \mu_D - \mu_M$.

On the other hand, from the zero profit cutoff condition (6) and the free entry condition (7), reductions in τ_X generate the following changes at the downstream stage:

$$\frac{d\ln\varphi_D}{d\ln\tau_X} = \frac{d\ln\varphi_{DM}}{d\ln\tau_X} < 0 < \frac{d\ln\varphi_X}{d\ln\tau_X} = \frac{d\ln\varphi_{XM}}{d\ln\tau_X}.$$

This means that, in contrast to input trade liberalization, resource reallocations arise across importing firms: more productive firms providing output to multiple markets expand by output trade liberalization, while less productive firms providing output in only a single market shrink by such liberalization. Despite this difference, however, firms that simultaneously export and import are more likely to benefit from output trade liberalization. While our results are similar to those in the existing literature exploring firms' export and import decisions, the disproportionate effect of trade liberalization on most globalized firms operates through different channels. For example, Bernard et al. (2018a) focus on strategic market power across a small number of such large firms where inputs are produced under conditions of perfect competition. In contrast, the present paper focuses on endogenous selection of measure-zero producers at the vertically-related stages where inputs are produced under conditions of imperfect competition.

Lemma 2: *If inputs and final goods are costly to trade across borders subject to selection under (16) and (18),*

(i) *Output trade liberalization gives rise to resource reallocations not only at the downstream stage but also at the upstream stage.*

(ii) *Output trade liberalization simultaneously induces such reallocations even across importing firms: more (less) productive importing firms providing output to multiple markets (a single market) expand (shrink).*

Our results in Lemmas 1 and 2 show that input trade liberalization may require less strict conditions than output trade liberalization in order to trigger resource reallocations at the vertically-related production stages: input trade liberalization requires only (16) while output trade liberalization requires both (16) and (18). More importantly, the difference in (15) and (17) illustrates potential channels through which input trade liberalization has a more significant impact on trade-induced productivity gains than output trade liberalization documented by empirical work (e.g., Amiti and Konings, 2007; Topalova and Khandelwal, 2011).[17] From the comparison of the impact on the domestic productivity cutoff at the *upstream* stage, it follows that

$$\left|\frac{d\ln\phi_D}{d\ln\tau_M}\right| > \left|\frac{d\ln\phi_D}{d\ln\tau_X}\right|,$$

and input trade liberalization always generates greater productivity gains relative to output trade liberalization. As for changes in the domestic productivity cutoff at the *downstream* stage,

$$\left|\frac{d\ln\varphi_D}{d\ln\tau_M}\right| > \left|\frac{d\ln\varphi_D}{d\ln\tau_X}\right| \quad \Longleftrightarrow \quad (\sigma-1-\theta)(\lambda_X-\lambda_M)+\theta(1-\mu_D+\mu_M)(1-\lambda_M) > 0.$$

Hence, the sufficient condition for this inequality is $\lambda_X \geq \lambda_M$, i.e., the domestic output share is greater than or equal to the domestic input share. This would be satisfied in current globalization where the input trade share is larger than the output trade share in the world trade volume.

[17]Following Melitz (2003), the aggregate productivity gains can be computed in terms of only the domestic productivity cutoffs.

3.2 Trade Flows

Having shown the impact of trade liberalization on resource reallocations at the two production stages, let us turn to the impact on trade flows. We continue to separately examine the impact of input trade liberalization and output trade liberalization, but note that trade liberalization in either type of goods affects trade flows of both types of goods. For example, input trade liberalization affects not only input trade flows directly but also output trade flows indirectly. In this section, thus, we first analyze the direct effect of trade liberalization on each type of goods, and then analyze the indirect effect of such liberalization.

Consider the impact of input trade liberalization on input trade flows. To see the sensitivity of such flows to changes in trade costs, we derive the (full) trade elasticity with respect to variable trade costs of inputs below. Following Melitz and Redding (2015) and using the domestic input share (9), this trade elasticity is given by

$$\begin{aligned}\varepsilon_M &= -\frac{d\ln\left(\frac{1-\lambda_M}{\lambda_M}\right)}{d\ln\tau_M} \\ &= \underbrace{(\sigma-1)}_{\text{Intensive margin elasticity}} + \underbrace{\left(-\frac{d\ln\Delta}{d\ln\tau_M}\right)}_{\substack{\text{Exporter extensive margin elasticity}\\ \text{in upstream stage}}} + \underbrace{\left(-\frac{d\ln\Lambda_M}{d\ln\tau_M}\right)}_{\substack{\text{Importer extensive margin elasticity}\\ \text{in downstream stage}}}.\end{aligned}$$

The fact that the extensive margin elasticity stems from the upstream and downstream stages indicates that reductions in variable trade costs of inputs allow not only suppliers to export their inputs more easily, but also firms to import inputs that are used in final goods production more easily. In other words, the presence of vertical linkages can amplify the effects of input trade costs on input trade flows, due to additional entry that takes place at another production stage. Furthermore, with a constant extensive margin elasticity $\theta_c = \vartheta_c = \theta$, the exporter (importer) extensive margin elasticity at the upstream (downstream) stage is also constant at

$$-\frac{d\ln\Delta}{d\ln\tau_M} = -\frac{d\ln\Lambda_M}{d\ln\tau_M} = \frac{\theta(\sigma-1)}{\sigma-1-\theta}.$$

As in changes in the domestic productivity cutoffs in (15), changes in the market shares associated with input trade liberalization are exactly the same between the upstream and downstream stages under assumption (14). Summing up the three terms, the input trade elasticity is expressed as

$$\varepsilon_M = \frac{(\sigma-1)(\sigma-1+\theta)}{\sigma-1-\theta}. \tag{19}$$

As for the impact of output trade liberalization on output trade flows, using the domestic output share (8), the (full) trade elasticity with respect to variable trade costs of final goods is given by

$$\begin{aligned}\varepsilon_X &= -\frac{d\ln\left(\frac{1-\lambda_X}{\lambda_X}\right)}{d\ln\tau_X} \\ &= \underbrace{(\sigma-1)}_{\text{Intensive margin elasticity}} + \underbrace{\left(-\frac{d\ln\Lambda_X}{d\ln\tau_X}\right)}_{\substack{\text{Exporter extensive margin elasticity}\\ \text{in downstream stage}}}.\end{aligned}$$

Since suppliers do not use final goods in input production in our model, output trade liberalization does not directly induce suppliers to enter and there is no importer extensive margin elasticity at the upstream stage.

This does not mean however that output trade liberalization has no impact on suppliers at the upstream stage. An expansion of exporting firms (by output trade liberalization) leads to an expansion of importing firms, which in turn leads to an expansion of exporting suppliers from changes in the market demands A, B. In other words, the presence of vertical linkages can also amplify the effects of output trade costs on output trade flows. Though the result seems natural, (17) shows that output trade liberalization triggers resource reallocations at the two production stages if and only if condition (18) holds, which relates to the extensive margin elasticity θ. In fact, the exporter extensive margin elasticity at the downstream stage is given by

$$-\frac{d \ln \Lambda_X}{d \ln \tau_X} = \theta \left[1 + \left(\frac{\sigma - 1 + \theta}{\sigma - 1 - \theta} \right) (\eta_D - \eta_X)(\mu_D - \mu_M) \right],$$

where η_D (η_X) is the domestic (foreign) input share of firms using only domestic input, satisfying $\eta_D - \eta_X > 0$ under (18).[18] Hence the output trade elasticity is

$$\varepsilon_X = \sigma - 1 + \theta \left[1 + \left(\frac{\sigma - 1 + \theta}{\sigma - 1 - \theta} \right) (\eta_D - \eta_X)(\mu_D - \mu_M) \right]. \tag{20}$$

As shown by Chaney (2008), the output trade elasticity can be decomposed into the intensive margin elasticity $\sigma - 1$ and the extensive margin elasticity θ when firm heterogeneity is present. While this decomposition applies, the extensive margin elasticity is greater in our two-stage production model than in the single-stage production model. The result can be seen from noting that the exporter extensive margin elasticity has an additional term which captures the feedback effect from an expansion of the upstream stage production to an expansion of the downstream stage production triggered by output trade liberalization. In this way, the impact of output trade liberalization on output trade flows can be amplified through vertical linkages, even though such liberalization does not directly induce suppliers to enter at the upstream stage.

Which trade elasticity is greater when both inputs and final goods are costly to trade subject to selection? Simple comparison of the trade elasticities in (19) and (20) immediately reveals that

$$\varepsilon_M > \varepsilon_X.$$

Thus the input trade elasticity is always greater than the output trade elasticity. This finding accords well with the widely known empirical fact that input trade has been growing faster than output trade in the real world (Hummels et al., 2001; Hanson et al., 2005; Johnson and Noguera, 2012).

The finding suggests that the gravity structure is drastically different between input trade and output trade. If ϕ and φ are distributed Pareto with a common scale parameter $\phi_{\min} = \varphi_{\min} = 1$ (keeping a common shape parameter γ), input trade flows $E_M = \int_v r_{Mv} dv$ and output trade flows $R_X = \int_i r_{Xi} di$ are decomposed into

$$E_M = \underbrace{\frac{\sigma(\sigma - 1 + \theta)}{\theta} k_M}_{\text{Average sales per supplier}} \times \underbrace{\left(\frac{1}{\phi_M} \right)^{\sigma - 1 + \theta} N_E}_{\text{Mass of suppliers}},$$

$$R_X = \underbrace{\frac{1}{\eta_X} \frac{\sigma(\sigma - 1 + \theta)}{\theta} f_X}_{\text{Average sales per firm}} \times \underbrace{\left(\frac{1}{\varphi_X} \right)^{\sigma - 1 + \theta} M_E}_{\text{Mass of firms}}.$$

[18] $\eta_D = \frac{1}{1 + \tau_M^{1-\sigma} \Delta \frac{V(\varphi_{DM})}{V(\varphi_D)}}$ and $\eta_X = \frac{1}{1 + \tau_M^{1-\sigma} \Delta \frac{V(\varphi_{XM})}{V(\varphi_X)}}$. Note that these always satisfy $1 > \eta_D - \eta_X$.

Moreover, the mass of entrants at each production stage is proportional to country size but independent of the productivity cutoffs under the distribution:

$$M_E = \frac{\sigma-1}{\sigma(\sigma-1+\theta)}\frac{L}{f_E}, \quad N_E = \frac{(\sigma-1)^2}{\sigma^2(\sigma-1+\theta)}\frac{L}{k_E}.$$

Substituting φ_X from (4) and ϕ_M from (10) as well as M_E, N_E derived above, we can express these trade flows as a gravity equation form:

$$\begin{aligned} E_M &= \psi_M \Lambda_M^{\frac{\sigma-1+\theta}{\sigma-1}} L A^{\frac{\sigma-1+\theta}{\sigma-1}} \tau_M^{-(\sigma-1+\theta)} k_M^{-\frac{\theta}{\sigma-1}}, \\ R_X &= \frac{\psi_X}{\eta_X} L B^{\frac{\sigma-1+\theta}{\sigma-1}} (\tau_X c_D)^{-(\sigma-1+\theta)} f_X^{-\frac{\theta}{\sigma-1}}, \end{aligned} \tag{21}$$

where ψ_M and ψ_X are some constant term.[19] As in a usual gravity equation, trade flows in either type of goods are a function of exporting country size L, importing country demands A, B, and bilateral trade barriers, both variable τ_M, τ_X and fixed k_M, f_X. The functional form is very similar to the gravity equation in Chaney (2008) in terms of the elasticity of trade flows with respect to trade barriers; however, input trade flows include the market share of importing firms Λ_M while output trade flows include the foreign input share of firms using only domestic inputs η_X, which help amplify the trade elasticity in the two-stage production model relative to that in the single-stage production model. For example, applying (14) to the market share of importing firms,

$$\Lambda_M^{\frac{\sigma-1+\theta}{\sigma-1}} = \left[\tau_M^{-(\sigma-1+\theta)} \left(\frac{\mu_D}{\mu_M}\right)^{\frac{\sigma-1}{\theta}} \left(\frac{f_{DM}}{f_D}\right)^{-1} \left(\frac{k_M}{k_D}\right)^{-\frac{\theta}{\sigma-1}}\right]^{\frac{\theta}{\sigma-1-\theta}}.$$

which is of course negatively affected by both variable and fixed trade costs. Simple inspection of E_M in (21) shows that the input trade costs τ_M decrease input trade flows not only through supplies' shipment with an elasticity of $\sigma - 1 + \theta (= \gamma)$ directly as in Chaney (2008), but also through firms' market share who make use of foreign inputs in production (i.e., through Λ_M) with an elasticity of $\frac{\theta(\sigma-1+\theta)}{\sigma-1-\theta}$ indirectly. The claim applies to output trade flows in the sense that they are indirectly affected through η_X.

To obtain the trade elasticity from the gravity equation in (21), it follows from the domestic shares λ_M, λ_X that the trade elasticities can be alternatively written as follows:[20]

$$\varepsilon_M = -\frac{d\ln(E_M/E_D)}{d\ln\tau_M}, \quad \varepsilon_X = -\frac{d\ln(R_X/R_D)}{d\ln\tau_X}.$$

Applying the Pareto distribution to the aggregate domestic expenditures $E_D = \int_v r_{Dv}dv$, $R_D = \int_i r_{Di}di$ yields the closed-form solutions of ε_M and ε_X which are exactly the same as those given in (19) and (20) respectively. Defining the trade elasticity with respect to fixed trade costs k_M, f_X similarly, this elasticity is also greater for input trade than for output trade. Hence, when estimating the elasticity of the value of trade with respect to trade barriers from the gravity equation, our model predicts that the trade elasticity of intermediate goods is *endogenously* greater than that of final goods.

The trade elasticity results are our first main proposition of the paper.

[19] $\psi_M = \frac{(\sigma-1)^2}{\sigma\theta k_E}$ and $\psi_X = \frac{\sigma-1}{\theta f_E}$. Note that (21) leads to the reduced-form expression of (1).

[20] This is similar to that in Arkolakis et al. (2012) who consider the "partial" trade elasticity focusing only on the direct effect of trade costs, while we consider the "full" trade elasticity taking account of all effects of trade costs. Although these two types of trade elasticity are generally different, they are the same under the Pareto distribution. See Melitz and Redding (2015).

Proposition 1: *If trade liberalization induces resource reallocations at the vertically-related production stages, the trade elasticity of intermediate goods is endogenously greater than that of final goods.*

It is worth stressing that our trade elasticity results do *not* come from the CES production function where final goods are produced by a variety of inputs. As is evident from above decomposition of the trade elasticity, the results come from the endogenous response in the extensive margin: reductions in trade costs allow new firms and suppliers to export and import at the respective production stage to different degrees. From this reasoning, it is possible to empirically test our theoretical predictions taking into account the different responses in the extensive margin. In a companion paper (Ara and Zhang, 2020), we study this channel by extending our model to a multi-sector, asymmetric-country setting and estimating the gravity equation derived from such extensions. Empirical support for Proposition 1 will be reviewed in Section 3.5.

We conclude this section by briefly mentioning the effect of output (input) trade barriers on input (output) trade flows. The trade elasticities capturing this indirect effect are defined as

$$\tilde{\varepsilon}_M = -\frac{d\ln\left(\frac{1-\lambda_M}{\lambda_M}\right)}{d\ln\tau_X}, \quad \tilde{\varepsilon}_X = -\frac{d\ln\left(\frac{1-\lambda_X}{\lambda_X}\right)}{d\ln\tau_M}.$$

It is clear to see that $\tilde{\varepsilon}_M$ and $\tilde{\varepsilon}_X$ have no intensive margin elasticity, meaning that output (input) trade barriers affect input (output) trade flows only through changes in the market share, i.e., extensive margin elasticity. Using the impact of trade liberalization on the productivity cutoffs in Lemmas 1 and 2, we get

$$\begin{aligned} \tilde{\varepsilon}_M &= (\sigma-1)\left(\frac{\sigma-1+\theta}{\sigma-1-\theta}\right)(\mu_D-\mu_M), \\ \tilde{\varepsilon}_X &= (\sigma-1)\left(\frac{\sigma-1+\theta}{\sigma-1-\theta}\right)(\eta_D-\eta_X). \end{aligned} \tag{22}$$

(22) shows that both elasticities have positive values if and only if there is complementarity between exporting and importing under (18). Consequently, input trade liberalization not only increases input trade flows directly, but also increases output trade flows indirectly, and vice versa.

3.3 Welfare Gains

Finally, we explore the welfare implications of trade liberalization in our model with vertical production. As in Sections 3.1 and 3.2, we separately study the impact of input trade liberalization and output trade liberalization on welfare in this section. Thus we focus on the welfare gains associated with trade liberalization in which trade costs of only one type of goods decrease, holding those of another type of goods fixed.

Totally differentiating (13), the changes in welfare are captured by the changes in the domestic productivity cutoffs at two production stages and those in the the domestic input share:

$$d\ln W = d\ln\phi_D + d\ln\varphi_D + \left(\frac{1}{\sigma-1}\right)d\ln\lambda_M. \tag{23}$$

The welfare changes in (23) show that not only does the domestic productivity cutoff at the downstream stage, but also the domestic productivity cutoff and the domestic input share at the upstream stage matter for welfare in our two-stage production model. This stands in a sharp contrast to the single-stage production model where the domestic productivity cutoff for firms is a sufficient statistic for welfare.

Consider the impact of input trade liberalization on welfare. On the one hand, the changes in the domestic productivity cutoffs associated with input trade liberalization are given by (15). Then, so long as (16) holds, (23) shows that reductions in such trade costs improve welfare by increasing the domestic productivity cutoffs at the two production stages. The welfare changes directly come from the Melitz-type resource reallocations among firms and suppliers in Lemma 1: input trade liberalization forces the least productive firms and suppliers to exit the respective production stage. Such two-sided selection amplifies the welfare gains from trade liberalization in our two-stage production model relative to those in the single-stage production model.[21] On the other hand, totally differentiating the domestic input share in (9) and noticing the definition of ε_M in Section 3.2, the changes in the domestic input share associated with input trade liberalization are given by

$$d \ln \lambda_M = (1 - \lambda_M)\varepsilon_M d \ln \tau_M.$$

While reductions in input trade costs naturally decrease the domestic input share, (23) shows that a decrease in the domestic input share *deteriorates* welfare. The reason for the counter-intuitive welfare changes is explained as follows. As seen in Section 2.4, welfare is defined as an inverse of the CES price index in the present model. Using φ_D in (4) and the final goods market demand B, this is given by

$$\frac{1}{P} = \left(\frac{\sigma - 1}{\sigma}\right)\left(\frac{L}{\sigma f_D}\right)^{\frac{1}{\sigma-1}} \frac{\varphi_D}{c_D}.$$

Thus, for given φ_D, welfare is negatively affected by the unit costs of firms using only domestic inputs c_D since the higher are the unit costs, the less efficient are these firms and the higher is the price index. Note that input trade liberalization endogenously affects the unit costs by altering the input availability to non-importing firms. Using ϕ_D in (10) and the input market demand A, the unit costs are expressed as

$$\frac{1}{c_D} = \left(\frac{\sigma - 1}{\sigma}\right)^{\frac{\sigma}{\sigma-1}} \left(\frac{L}{\sigma k_D}\right)^{\frac{1}{\sigma-1}} \phi_D \lambda_M^{\frac{1}{\sigma-1}}.$$

Thus, for given ϕ_D, the unit costs are negatively affected by the domestic input share λ_M since the lower is this share, the less available are domestic inputs for non-importing firms and the higher is the price index.[22] This means that when there is selection into importing among firms, input trade liberalization can have a negative impact on welfare by increasing the unit costs of firms who cannot access foreign inputs. However, this negative effect is always dominated by the positive effect from selection. Summing up the three terms in (23) where ε_M in $d \ln \lambda_M$ is given in (19), the welfare changes associated with input trade liberalization are simply given by

$$d \ln W = -(1 - \lambda_M) d \ln \tau_M, \tag{24}$$

which indicates that input trade liberalization is always welfare-enhancing. More important is that the elasticity of welfare with respect to input trade costs is equivalent to the foreign input share $1 - \lambda_M$.

Consider next the impact of output trade liberalization on welfare. The changes in the domestic productivity cutoffs associated with output trade liberalization are given by (17). Then, so long as (16) and (18) hold, (23) shows that reductions in such trade costs improve welfare by increasing the domestic productivity cutoffs at

[21] The welfare results in vertical production bear some resemblance to those in sequential production (Melitz and Redding, 2014), in the sense that when final goods are produced from traded intermediate goods, the welfare gains are amplified through an increase in domestic productivity. The key difference is that each stage is imperfectly competitive along with selection in our setting.

[22] Substituting $1/c_D$ into $1/P$ above, we get welfare per worker in (13).

the two production stages. The welfare changes again directly come from the Melitz-type resource reallocations among firms and suppliers in Lemma 2. As for the impact on the domestic input share, in contrast, output trade liberalization indirectly increases input trade flows and hence decreases the domestic input share, due to complementarity between exporting and importing under (18). Noting that the corresponding trade elasticity is $\tilde{\varepsilon}_M$ in (22), the changes in the domestic input share associated with output trade liberalization are given by

$$d\ln\lambda_M = (1-\lambda_M)\tilde{\varepsilon}_M d\ln\tau_X.$$

Then (23) shows that reductions in τ_X deteriorate welfare, which is explained by the impact on the unit costs of non-importing firms. While each channel of the welfare changes is qualitatively similar between input and output trade liberalization, the overall welfare changes are quantitatively different between these trade shocks, as two-sided reallocations take place on a different scale (see (15) and (17)), which in turn generates a different impact on the domestic input share. Summing up the three terms in (23), the welfare changes associated with output trade liberalization are simply given by

$$d\ln W = -(1-\lambda_X)d\ln\tau_X. \tag{25}$$

It is important to observe that the elasticity of welfare with respect to output trade costs is equivalent to the foreign output share $1-\lambda_X$.

We are now ready for comparing the welfare changes associated with input and output trade liberalization. It follows immediately from (24) and (25) that

$$\left|\frac{d\ln W}{d\ln\tau_M}\right| > \left|\frac{d\ln W}{d\ln\tau_X}\right| \quad \Longleftrightarrow \quad \lambda_X > \lambda_M.$$

Thus, the welfare gains from input trade liberalization are greater than those from output trade liberalization if and only if the domestic output share is greater than the domestic input share. Recall that this is the sufficient condition under which input trade liberalization induces the greater changes in the domestic productivity cutoffs at the downstream stage relative to output trade liberalization. In this sense, the welfare results are consistent with the resource-reallocation results in Lemmas 1 and 2.

We conclude this section by relating our welfare results to the Arkolakis et al. (2012) welfare formula. Using the relationship between the changes in the domestic input share and those in input trade costs, we can express the welfare changes associated with input trade liberalization (24) as

$$d\ln W = -\frac{d\ln\lambda_M}{\varepsilon_M}.$$

Hence, the welfare changes associated with input trade liberalization can be captured by only the two sufficient statistics: the domestic input share λ_M and the input trade elasticity ε_M, even with two-sided heterogeneity. Similarly, totally differentiating the domestic output share in (8) and noting the definition of ε_X, the impact of output trade costs on the domestic output share is given by $d\ln\lambda_X = (1-\lambda_X)\varepsilon_X d\ln\tau_X$. Using this relationship, the welfare changes associated with output trade liberalization (25) can be expressed in terms of the domestic output share λ_X and the output trade elasticity ε_X only:

$$d\ln W = -\frac{d\ln\lambda_X}{\varepsilon_X}.$$

Integrating the expressions between the initial equilibrium (before trade liberalization) and the new equilibrium (after trade liberalization), the welfare changes associated with trade liberalization are given by (2). Therefore, the Arkolakis et al. (2012) welfare formula applies here: the welfare changes can be computed only from the two statistics, λ_j and ε_j, for each type of goods j, which means that the welfare gains are the same between these different trade shocks conditional on these statistics.[23] More importantly, our welfare comparison between input and output trade liberalization illustrates the main emphasis of Arkolakis et al. (2012) in a very clear manner. As we have seen at the end of Section 3.1, the productivity gains are greater for input trade liberalization than for output trade liberalization holding the two kinds of domestic share equal. Given that, it is quite natural to imagine that the welfare gains are also greater for input trade liberalization than for output trade liberalization. Our analysis demonstrates that this is not the case. In our model, the difference in two trade elasticities reflects that there is an extra adjustment in the extensive margin in input trade that is absent in output trade, which can amplify the welfare gains from input trade liberalization. Yet, conditional on the two statistics, this extra margin only affects the composition of the welfare gains from input trade liberalization, not their total size.

Though useful, (2) implies that the welfare changes associated with input and output trade liberalization are the same, holding the two kinds of domestic share equal. There is however mounting evidence suggesting that input trade has been growing faster than output trade and its share is increasingly greater in the world trade volume. These pieces of evidence indicate that the welfare evaluation under such conditioning might lead to inaccurate understanding of globalization where fragmentation of production processes plays a prominent role in improving welfare in each country. The finding that the welfare gains from trade liberalization are greater for input trade than for output trade under rapidly rising input trade is also consistent with empirical evidence that input tariff reductions increase industry productivity more than output tariff reductions, because such productivity improvements are typically associated with the welfare gains from input trade liberalization more than from output trade liberalization.

The welfare results are our second main proposition of the paper.

Proposition 2: *If trade liberalization induces resource reallocations at the vertically-related production stages, the welfare gains from input trade liberalization are greater than those from output trade liberalization if and only if the domestic input share is smaller than the domestic output share.*

3.4 Extensions

The main analysis has assumed a common Pareto shape parameter shared by upstream and downstream stages. Recent empirical work, however, finds that there could be substantial differences in the degree of heterogeneity between these stages (Bernard et al., 2022) in which case the differences in the productivity distribution between firms and suppliers could have an influence on the trade elasticity results. For example, if the upstream stage is more homogeneous than the downstream stage, suppliers are more sensitive to changes in trade costs than firms, offering an additional channel through which to increase the input trade elasticity relative to the output trade elasticity. This new channel may alter the welfare results as well. In that sense, incorporating differences in the degree of heterogeneity between stages is key to addressing the extent to which our results can be generalized. This section briefly argues that the main results continue to hold even with a different Pareto shape parameter. Detailed analyses are available in the online Appendix (Section B.3).

[23] The reason why the changes in welfare cannot be expressed in terms of the changes in domestic share of both input and output is explained by the fact that we have two separate welfare exercises, i.e., reductions in input trade costs and output trade costs. Consequently, it is not straightforward to condense down these welfare changes to a single formula that is easy to interpret.

When the Pareto shape parameter differs by stage, the extensive margin elasticity also differs by stage, i.e., $\theta \neq \vartheta$ and assumption (14) no longer holds. This gives rise to the following new results.

First, changes in the domestic productivity cutoffs associated with input trade liberalization are different between the two production stages, which are the same in the baseline case (see (15)). Specically, changes in φ_D are greater than those in ϕ_D if and only if $\theta > \vartheta$, implying when the upstream stage is more homogeneous than the downstream stage, input trade liberalization has a greater impact on firms than on suppliers. Intuitively, reductions in input trade costs allow suppliers to export their inputs relatively more easily in this general case than in the baseline case. As the availability of foreign inputs improves at the downstream stage, competition among firms becomes tougher via lower unit costs, which increases φ_D relatively more than ϕ_D. In contrast, changes in the domestic productivity cutoffs associated with output trade liberalization are different between the two production stages as in the baseline case (see (17)). Comparing changes in ϕ_D and φ_D associated with these different trade shocks, we find that input trade liberalization still generates a more significant impact on productivity gains than output trade liberalization. Thus, while the impact on the productivity cutoffs requires some qualification, the results in Lemmas 1 and 2 hold even with a different shape parameter.

Second, the trade elasticity includes a different Pareto shape parameter between the two production stages. Using changes in the productivity cutoffs above, the trade elasticities in (19) and (20) are given by

$$\varepsilon_M = \frac{(\sigma-1)(\sigma-1+\theta)(\sigma-1+\vartheta)}{(\sigma-1)^2-\theta\vartheta},$$

$$\varepsilon_X = \sigma-1+\vartheta+\theta\left(\frac{(\sigma-1+\vartheta)^2}{(\sigma-1)^2-\theta\vartheta}\right)(\eta_D-\eta_X)(\mu_D-\mu_M),$$

where $(\sigma-1)^2-\theta\vartheta > 0$. Not surprisingly, if $\theta = \vartheta$, these expressions collapse to those in the baseline model. Further, we can see that both ε_M and ε_X are increasing in θ and thus a homogeneous upstream stage leads to an increase in both input trade flows and output trade flows. Consider the reason why the input trade elasticity ε_M is increasing in θ: as θ is larger, suppliers are more sensitive to changes in trade costs and reductions in input trade costs allow more suppliers to start exporting at the upstream stage; at the same time, this increase in input trade flows also improves the availability of foreign inputs, allowing more firms to start importing at the downstream stage. This channel operating through the exporter and importer extensive margin elasticities amplifies the input trade elasticity when the upstream stage is more homogeneous than the downstream stage. The reason why the output trade elasticity ε_X is also increasing in θ is similarly explained: as θ is larger, input trade flows are greater, allowing more firms to start exporting at the downstream stage via the love-of-variety effect. Interestingly, the closed-form solution of ε_X shows that the output trade elasticity is decomposed into the intensive margin elasticity $\sigma-1$, the extensive margin elasticity at the downstream stage ϑ and the extensive margin elasticity at the upstream stage θ weighted by a composite term that is positive if and only if (18) holds. Despite this new outcome in the extensive margin, simple inspection of the trade elasticities immediately reveals that the input trade elasticity is still greater than the output trade elasticity for any combination of the Pareto shape parameters, and hence the trade elasticity results in Proposition 1 hold in this general case.

Finally, changes in welfare associated with trade liberalization need to be modified so as to reflect qualification of changes in the productivity cutoffs seen above. Though welfare changes include both θ and ϑ, we find that these terms are exactly cancelled out by one another and changes in welfare associated with input and output trade liberalization are simply given by (24) and (25), respectively. As a consequence, the elasticity of welfare with respect to input (output) trade costs is still equivalent to the foreign input (output) share. Thus, we have the same welfare results presented in Proposition 2, which would be quite special to the Pareto distribution.

3.5 Evidence

In this section, we provide evidence supporting our trade elasticity results, drawing upon our companion paper (Ara and Zhang, 2020). The authors extend our setting to a multi-sector, asymmetric-country setting to allow for potential differences across countries; however complexity of the model forces them to impose a restriction that either inputs or final goods are only tradable.[24] Nevertheless, the finding in Proposition 1 is qualitatively the same, and the current model setup offers a more comprehensive theoretical framework that can be empirically investigated in a similar way, albeit at the expense of resorting to a symmetric-country setting.

Let us first show the empirical specification and hypotheses from our theory. When inputs and final goods are costly to trade across borders subject to selection, the gravity equation of each type of goods is given as (21) whereas its reduced-form expression is given as (1). Applying a log-linear approximation to (1) leads to the following specification for the total value of trade flows between country A and country B for goods $j \in \{M, X\}$:

$$\ln T_j^{AB} = \delta^A + \delta^B - \varepsilon_1 \ln \tau_j^{AB} - \varepsilon_2 D_j \ln \tau_j^{AB} + \delta_j^{AB},$$

where δ^A is country A's fixed effect, δ^B is country B's fixed effect, D_j is a dummy variable that equals one (zero) if the observation is categorized as intermediate (final) goods, and δ_j^{AB} is an orthogonal error term. While we report the estimation results when variable trade costs are measured by distance (which is common between inputs and final goods), the results similarly hold when such costs are measured by tariffs (which are different between the two types of goods).[25] As the specification indicates, the coefficients ε_1 and $\varepsilon_1 + \varepsilon_2$ are the trade elasticity for intermediate goods and final goods, respectively, so that $\varepsilon_1 + \varepsilon_2 = \varepsilon_M$ in (19) and $\varepsilon_1 = \varepsilon_X$ in (20). Armed with this specification, we can test the following two hypotheses that directly come from Proposition 1. First, the trade elasticity is greater for intermediate goods than for final goods. Second, the difference in the trade elasticity comes mainly from the extensive margin elasticity rather than the intensive margin elasticity. The latter is tested by decomposing the total value of trade flows as $T_j^{AB} = M_j^{AB} \times \bar{T}_j^{AB}$ where M_j^{AB} is the number of producers (extensive margin) while $\bar{T}_j^{AB}$ is the average trade value per producer (intensive margin), as we have shown in deriving the gravity equation (21).

In assessing these patterns, the authors divide China's imports in terms of US dollar into intermediate goods and final goods applying the UN Broad Economic Categories classification to the China Customs database (at the 6-digit HS product level) in 2000–2007. Given that this was a period of drastic trade liberalization for both intermediate goods and final goods that took place after China's join to WTO, the data setting is particularly apt for performing this test as such liberalization is expected to have a critical impact on the extensive margin as well as the intensive margin. For each product imported from each trading partner in each year, the total value of imports of two types of goods is decomposed into the number of importing firms with positive trade flows (extensive margin) and the average import value conditional on positive trade flows (intensive margin). Using the empirical specification above, they separately estimate the gravity equation of intermediate goods and final goods to see how the trade elasticity with respect to variable trade costs differs across these two types of goods.

[24] To see the source of complexity, recall that we have solved for the general equilibrium of our model characterized by the eight unknowns. If the present model is extended to two asymmetric countries where the relative wage is fixed by an outside good sector, the number of unknowns is doubled and it is difficult to obtain analytical solutions. Avoiding this difficulty, Ara and Zhang (2020) solve the model by assuming that variable trade costs of either inputs or final goods are prohibitively high, which enables them to analytically characterize equilibrium. Naturally the closed-form solutions of trade elasticity differ from (19) and (20) in this setting, but the empirical specification and testable hypotheses are similar to those described below.

[25] Ara and Zhang (2020) first consider distance as the measure of variable trade costs and interpret the distance elasticity as the trade elasticity. Given that the authors focus on China's imports only, they cannot control for the source and destination country fixed effects when estimating the distance elasticity. To see robustness of the results, they then consider tariffs as the measure of variable trade costs and find qualitatively similar results for the tariff elasticity by controlling for the source country fixed effects.

Table 1 – Estimates of overall goods, intermediate goods, and final goods

	Overall			Intermediate			Final		
	Total	Extensive	Intensive	Total	Extensive	Intensive	Total	Extensive	Intensive
Distance	0.758	0.549	0.208	0.797	0.574	0.223	0.689	0.508	0.181
	(0.015)	(0.007)	(0.010)	(0.020)	(0.009)	(0.013)	(0.023)	(0.010)	(0.016)
No. of obs	576,509	576,509	576,509	354,976	354,976	354,976	220,693	220,693	220,693
Adj. R^2	0.403	0.498	0.389	0.372	0.510	0.343	0.443	0.478	0.446

Source: Ara and Zhang (2020, Tables 3 and 4)

Notes: Standard errors clustered at product-level are in parentheses. Product and year fixed effects are included, while source and destination country fixed effects are not included as the authors consider the distance elasticity focusing on China's imports only. All results are statistically significant at the 1% level.

Table 2 – Estimates with an input dummy interaction

	Total	Extensive	Intensive
Distance	0.718^{***}	0.516^{***}	0.202^{***}
	(0.022)	(0.010)	(0.015)
Distance $*$ Input	0.064^{**}	0.054^{***}	0.010
	(0.029)	(0.013)	(0.019)
No. of obs	575,669	575,669	575,669
Adj. R^2	0.403	0.498	0.388

Source: Ara and Zhang (2020, Table 5)

Notes: Standard errors clustered at product-level are in parentheses. Product and year fixed effects are included, while source and destination country fixed effects are not included as the authors consider the distance elasticity focusing on China's imports only. ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively.

In order to make reference to the existing literature, the authors also estimate the gravity equation of overall goods without distinction of the two types of goods. Finally, they run the regression with an input dummy interaction and examine whether there is a statistically significant difference in the trade elasticity between the two types of goods. In this regression, they pool their dataset on intermediate goods and final goods together and check the coefficients on variable trade costs for intermediate goods relative to final goods.

Table 1 reports the estimates for the trade elasticity that is decomposed into the extensive and intensive margins for each type of goods where all variables are measured in logs. In the estimates of overall goods, the negative relationship between total import value and distance is largely accounted for by the extensive margin, which accords well with previous work that studies the role of the extensive and intensive margins in the gravity equation (e.g., Bernard et al., 2007). When comparing the estimates between the two types of goods, however, they find that the coefficient of distance on the total import value is greater for intermediate goods than for final goods, while retaining the major role of the extensive margin in the gravity for both types of goods. Table 2, on the other hand, reports the estimates that include the input dummy interaction. The coefficient of distance on the interaction term for the total import value shows that the trade elasticity is significantly greater for intermediate goods than for final goods. Looking at the coefficient of distance on the interaction term for the extensive and intensive margins further shows that the difference in the trade elasticity between the two types of goods is significantly explained only through the extensive margin. These pieces of evidence support a new prediction of our model: the trade elasticity is significantly greater for intermediate goods than for final goods due mainly to the extensive margin.

4 Conclusion

This paper has presented a heterogeneous firm model to analyze selection effects at different production stages on trade-induced intra-industry resource reallocations. We show that the trade elasticity of intermediate goods is endogenously greater than that of final goods, due to an extra adjustment in the extensive margin. We also find that the welfare gains from input trade liberalization are greater than those from output trade liberalization if and only if the domestic input share is smaller than the domestic output share. These findings could help us to obtain a better understanding about rapidly rising input trade and large productivity gains associated with input trade liberalization reported by empirical work. One of the broader policy implications from our analysis is that, when both intermediate goods and final goods are costly to trade across borders subject to selection, the difference in the trade elasticity between these two types of goods is crucial to understanding the mechanism that generates the difference in the welfare gains from trade liberalization.

To highlight an extra adjustment of the extensive margin in input trade that is absent in final goods trade, we have resorted to a two-country symmetric setting. While our model can be easily extended to a multi-country symmetric setting, it is challenging to develop a multi-country asymmetric setting in which trade costs have a potentially asymmetric impact on input and output trade flows across trading partners. We expect that this setting could provide an important channel through which to reinforce the trade elasticity results via a positive correlation between the productivity level of firms and the number of source countries. It is also interesting to investigate the impact of country asymmetry on trade patterns between intermediate goods and final goods as well as general-equilibrium consequences of such specialization patterns for welfare gains. Does a larger (smaller) country host disproportionally more firms (suppliers) and become a net exporter of final (intermediate) goods in vertical production? Does input trade liberalization induce agglomeration of suppliers (firms) in a liberalizing (non-liberalizing) country and give rise to different welfare gains from those from output trade liberalization? We leave these questions to future work.

A Proofs

A.1 Proof of Lemma 1

To show the derivation of (15), we first derive changes by the variable trade costs of inputs at the downstream stage. Taking the log and differentiating the zero profit cutoff conditions (5) and (6) with respect to τ_M,

$$\begin{aligned} d\ln\varphi_X - d\ln\varphi_D &= 0, \\ d\ln\varphi_{XM} - d\ln\varphi_{DM} &= 0, \\ d\ln\varphi_{DM} - d\ln\varphi_D &= -\frac{1}{\sigma-1}d\ln\Delta + d\ln\tau_M, \\ d\ln\varphi_{XM} - d\ln\varphi_X &= -\frac{1}{\sigma-1}d\ln\Delta + d\ln\tau_M. \end{aligned} \tag{A.1}$$

Using the definition of Δ and $\theta_c = \theta$,

$$d\ln\Delta = -\theta(d\ln\phi_M - d\ln\phi_D).$$

Moreover, differentiating the free entry condition (7) with respect to τ_M,

$$\sum_c f_c J'(\varphi_c)\varphi_c d\ln\varphi_c = 0.$$

Solving this for $d\ln\varphi_{DM}$ and $d\ln\varphi_{XM}$ by using $d\ln\varphi_X = d\ln\varphi_D$ and $d\ln\varphi_{XM} = d\ln\varphi_{DM}$ from (A.1),

$$d\ln\varphi_{DM} = -\alpha d\ln\varphi_D, \tag{A.2}$$

where

$$\alpha \equiv \frac{f_D J'(\varphi_D)\varphi_D + f_X J'(\varphi_X)\varphi_X}{f_{DM}J'(\varphi_{DM})\varphi_{DM} + f_{XM}J'(\varphi_{XM})\varphi_{XM}}.$$

Just like (4) and (7) cannot characterize the levels at the downstream stage, (A.1) and (A.2) cannot characterize the changes at the downstream stage through the changes in the market share $d\ln\Delta$.

Next we calculate changes by the variable trade costs of inputs at the upstream stage. Taking the log and differentiating the zero profit cutoff condition (11) with respect to τ_M,

$$d\ln\phi_M - d\ln\phi_D = -\frac{1}{\sigma-1}d\ln\Lambda_M + d\ln\tau_M. \tag{A.3}$$

Using the definition of Λ_M and $\vartheta_c = \theta$ as well as $d\ln\varphi_X = d\ln\varphi_D$ and $d\ln\varphi_{XM} = d\ln\varphi_{DM}$ from (A.1),

$$d\ln\Lambda_M = -\theta(d\ln\varphi_{DM} - d\ln\varphi_D).$$

Moreover, differentiating the free entry condition (12) with respect to τ_M,

$$\sum_c k_c J'(\phi_c)\phi_c d\ln\phi_c = 0.$$

Solving this for $d\ln\phi_M$, we have

$$d\ln\phi_M = -\beta d\ln\phi_D, \tag{A.4}$$

where

$$\beta \equiv \frac{k_D J'(\phi_D)\phi_D}{k_M J'(\phi_M)\phi_M}.$$

Note the similarity between (A.1) and (A.3) as well as (A.2) and (A.4).

Finally, we solve for the changes in the economy by the variable trade costs of inputs, taking into account vertical linkages between the two production stages. Substituting (A.2) and (A.4) into the third equation of (A.1), and substituting (A.2) and (A.4) into (A.3) respectively,

$$\begin{aligned} -(\alpha+1)d\ln\varphi_D &= -\frac{\theta}{\sigma-1}(\beta+1)d\ln\phi_D + d\ln\tau_M, \\ -(\beta+1)d\ln\phi_D &= -\frac{\theta}{\sigma-1}(\alpha+1)d\ln\varphi_D + d\ln\tau_M. \end{aligned}$$

Solving these for $d\ln\phi_D$ and $d\ln\varphi_D$ yields

$$\begin{aligned} d\ln\phi_D &= -\left(\frac{\sigma-1}{(\sigma-1-\theta)(\beta+1)}\right)d\ln\tau_M, \\ d\ln\varphi_D &= -\left(\frac{\sigma-1}{(\sigma-1-\theta)(\alpha+1)}\right)d\ln\tau_M. \end{aligned}$$

It remains to show $\frac{1}{\alpha+1} = \frac{1}{\beta+1} = 1-\lambda_M$. Differentiating $J(\varphi_c) = \int_{\varphi_c}^{\infty}\left[\left(\frac{\varphi}{\varphi_c}\right)^{\sigma-1} - 1\right]dG(\varphi)$ with respect to φ_c,

$$J'(\varphi_c) = -\left(\frac{\sigma-1}{\varphi_c}\right)[J(\varphi_c) + 1 - G(\varphi_c)].$$

From $V(\varphi_c) = \int_{\varphi_c}^{\infty}\varphi^{\sigma-1}dG(\varphi)$, it follows that $J(\varphi_c) + 1 - G(\varphi_c) = \varphi_c^{1-\sigma}V(\varphi_c)$ and hence

$$J'(\varphi_c) = -(\sigma-1)\varphi_c^{-\sigma}V(\varphi_c), \tag{A.5}$$

which also holds for suppliers' productivity cutoff ϕ_c. Substituting (A.5) into the definition of α and β above, and subsequently using the zero profit cutoff conditions (4) and (10),

$$\alpha = \beta = \frac{1}{\tau_M^{1-\sigma}\Delta\Lambda_M}. \tag{A.6}$$

The result follows immediately from using (A.6) in the definition of λ_M.

A.2 Proof of Lemma 2

To show the derivation of (17), we closely follow the steps in Appendix A.1. Taking the log and differentiating the zero profit cutoff condition (5) and (6) with respect to τ_X,

$$\begin{aligned} d\ln\varphi_X - d\ln\varphi_D &= d\ln\tau_X, \\ d\ln\varphi_{XM} - d\ln\varphi_{DM} &= d\ln\tau_X, \\ d\ln\varphi_{DM} - d\ln\varphi_D &= -\frac{1}{\sigma-1}d\ln\Delta, \\ d\ln\varphi_{XM} - d\ln\varphi_X &= -\frac{1}{\sigma-1}d\ln\Delta, \end{aligned} \tag{A.7}$$

where $d\ln\Delta$ is the same expression in Appendix A1. Moreover, differentiating the free entry condition (7) with respect to τ_X also yields the same expression in Appendix A.1. Solving this for $d\ln\varphi_{DM}$ and $d\ln\varphi_{XM}$ by noting $d\ln\varphi_X = d\ln\varphi_D + d\ln\tau_X$ and $d\ln\varphi_{XM} = d\ln\varphi_{DM} + d\ln\tau_X$ in (A.7),

$$d\ln\varphi_{DM} = -\alpha d\ln\varphi_D - \tilde{\alpha} d\ln\tau_X, \tag{A.8}$$

where

$$\tilde{\alpha} \equiv \frac{f_X J'(\varphi_X)\varphi_X + f_{XM} J'(\varphi_{XM})\varphi_{XM}}{f_{DM} J'(\varphi_{DM})\varphi_{DM} + f_{XM} J'(\varphi_{XM})\varphi_{XM}}.$$

As for the equilibrium in changes at the upstream stage, taking the log and differentiating the zero profit cutoff condition (11) with respect to τ_X,

$$d\ln\phi_M - d\ln\phi_D = -\frac{1}{\sigma-1} d\ln\Lambda_M, \tag{A.9}$$

where

$$d\ln\Lambda_M = -\theta(d\ln\varphi_{DM} - d\ln\varphi_D) - (\sigma - 1 + \theta)(\mu_D - \mu_M) d\ln\tau_X.$$

Moreover, differentiating the free entry condition (12) with respect to τ_X yields

$$d\ln\phi_M = -\beta d\ln\phi_D. \tag{A.10}$$

Finally, substituting (A.8) and (A.10) into the third equation of (A.7), and substituting (A.8) and (A.10) into (A.9) respectively,

$$\begin{aligned}
-(\alpha+1) d\ln\varphi_D &= -\frac{\theta}{\sigma-1}(\beta+1) d\ln\phi_D + \tilde{\alpha} d\ln\tau_X, \\
-(\beta+1) d\ln\phi_D &= -\frac{\theta}{\sigma-1}(\alpha+1) d\ln\varphi_D - \left(\frac{\theta\tilde{\alpha} - (\sigma-1+\theta)(\mu_D-\mu_M)}{\sigma-1}\right) d\ln\tau_X.
\end{aligned}$$

Solving for $d\ln\phi_D$ and $d\ln\varphi_D$ yields

$$\begin{aligned}
d\ln\phi_D &= -\left(\frac{(\sigma-1)(\mu_D-\mu_M)}{(\sigma-1-\theta)(\beta+1)}\right) d\ln\tau_X, \\
d\ln\varphi_D &= -\left(\frac{\tilde{\alpha}(\sigma-1-\theta) + \theta(\mu_D-\mu_M)}{(\sigma-1-\theta)(\alpha+1)}\right) d\ln\tau_X.
\end{aligned}$$

Moreover, using (A.5) into the definition of $\tilde{\alpha}$ above,

$$\tilde{\alpha} = (1-\lambda_X)(\alpha+1).$$

The result follows immediately from using this and (A.6) in the definition of λ_M.

References

Amiti M, Konings J. 2007. Trade Liberalization, Intermediate Inputs, and Productivity. *American Economic Review* 97, 1611-1638.

Antràs P, Chor D. 2022. Global Value Chains. *Handbook of International Economics* Volume 5, Elsevier: North Holland, 297-376.

Antràs P, de Gortari A. 2020. On the Geography of Global Value Chains. *Econometrica* 88, 1553-1598.

Antràs P, Fort TC, Tintelnot F. 2017. The Margins of Global Sourcing: Theory and Evidence from US Firms. *American Economic Review* 107, 2514-2564.

Antràs P, Helpman E. 2004. Global Sourcing. *Journal of Political Economy* 112, 552-580.

Ara T, Ghosh A. 2016. Tariffs, Vertical Specialization and Oligopoly. *European Economic Review* 82, 1-23.

Ara T, Zhang H. 2020. The Margins of Intermediate Goods Trade: Theory and Evidence. *The International Economy* 23, 105-144.

Arkolakis C, Costinot A, Rodriguez-Clare A. 2012. New Trade Models, Same Old Gains? *American Economic Review* 102, 94-130.

Bas M, Mayer T, Thoenig M. 2017. From Micro to Macro: Demand, Supply, and Heterogeneity in the Trade Elasticity. *Journal of International Economics* 108, 1-19.

Bernard AB, Dhyne E, Magerman G, Manova K. 2022. The Origins of Firm Heterogeneity: A Production Network Approach. *Journal of Political Economy* 130, 1765-1804.

Bernard AB, Jensen JB, Redding SJ, Schott SK. 2007. Firms in International Trade. *Journal of Economic Perspectives* 21, 105-130.

Bernard AB, Jensen JB, Redding SJ, Schott SK. 2012. The Empirics of Firm Heterogeneity and International Trade. *Annual Review of Economics* 4, 283-313.

Bernard AB, Jensen JB, Redding SJ, Schott SK. 2018a. Global Firms. *Journal of Economic Literature* 56, 565-619.

Bernard AB, Moxnes A, Ulltveit-Moe KH. 2018b. Two-sided Heterogeneity and Trade. *Review of Economics and Statistics* 100, 424-439.

Blaum J. 2019. Global Firms in Large Devaluations. Mimeo.

Caliendo L, Feenstra RC, Romalis J, Taylor AM. 2021. A Second-best Argument for Low Optimal Tariffs. Mimeo.

Chaney T. 2008. Distorted Gravity: The Intensive and Extensive Margins of International Trade. *American Economic Review* 98, 1707-1721.

Costinot A, Rodríguez-Clare A. 2014. Trade Theory with Numbers: Quantifying the Consequences of Globalization. *Handbook of International Economics* Volume 4, Elsevier: North Holland, 197-261.

Eaton J, Kortum S. 2002. Technology, Geography, and Trade. *Econometrica* 70, 1741-1779.

Ghosh A, Morita H. 2007. Free Entry and Social Efficiency under Vertical Oligopoly. *RAND Journal of Economics* 38, 541-554.

Goldberg PK, Khandelwal AK, Pavcnik N, Topalova P. 2010. Imported Intermediate Inputs and Domestic Product Growth: Evidence from India. *Quarterly Journal of Economics* 125, 1727-1767.

Gopinath G, Neiman B. 2014. Trade Adjustment and Productivity in Large Crises. *American Economic Review* 104, 793-831.

Halpern L, Koren M, Szeidl A. 2015. Imported Inputs and Productivity. *American Economic Review* 105, 3660-3703.

Hanson GH, Mataloni RJ, Slaughter MJ. 2005. Vertical Production Networks in Multinational Firms. *Review of Economics and Statistics* 87, 664-678.

Hummels D, Ishii J, Yi KM. 2001. The Nature and Growth of Vertical Specialization in World Trade. *Journal of International Economics* 54, 75-96.

Ishikawa J, Spencer BJ. 1999. Rent-shifting Export Subsidies with an Imported Intermediate Product. *Journal of International Economics* 48, 199-232.

Johnson RC, Moxnes A. 2019. GVCs and Trade Elasticities with Multistage Production. Mimeo.

Johnson RC, Noguera G. 2012. Accounting for Intermediates: Production Sharing and Trade in Value Added. *Journal of International Economics* 86, 224-236.

Kasahara H, Lapham B. 2013. Productivity and the Decision to Import and Export: Theory and Evidence. *Journal of International Economics* 89, 297-316.

Melitz MJ. 2003. The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity. *Econometrica* 71, 1695-1725.

Melitz MJ, Redding SJ. 2014. Missing Gains from Trade? *American Economic Review* 104, 317-321.

Melitz MJ, Redding SJ. 2015. New Trade Models, New Welfare Implications. *American Economic Review* 105, 1105-1146.

Ossa R. 2015. Why Trade Matters After All. *Journal of International Economics* 97, 266-277.

Topalova P, Khandelwal A. 2011. Trade Liberalization and Firm Productivity: The Case of India. *Review of Economics and Statistics* 93, 995-1009.

Online Appendix (Not for Publication)

B.1 Derivations

This section contains detailed derivations for the expressions omitted in the main text due to the space constraint. The same section names are attached to the following sections where the detailed derivations are required.

B.1.1 Firms

To derive firm i's input expenditure e_i, substituting x_{Div} and x_{Miv} into the CES production function, we get $x_i = e_i/c_i$. Rewriting this as $e_i = c_i x_i$ and using firm i's technology $q_i = \varphi x_i$ yields $e_i = c_i q_i/\varphi$.

To derive firm i's unit costs in (3), note that $c_D^{1-\sigma} = \int_v p_{Dv}^{1-\sigma} dv$ is expressed as

$$c_D^{1-\sigma} = N_E \int_{\phi_D}^{\infty} p_{Dv}^{1-\sigma} dG(\phi).$$

Using $p_{Dv} = \frac{\sigma}{\sigma-1}\frac{1}{\phi}$ and the definition of $V(\phi_c)$, we get

$$c_D^{1-\sigma} = N_E \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} V(\phi_D).$$

Similarly, $c_M^{1-\sigma} = \int_v p_{Mv}^{1-\sigma} dv$ is expressed as

$$c_M^{1-\sigma} = N_E \left(\frac{\sigma\tau_M}{\sigma-1}\right)^{1-\sigma} V(\phi_M).$$

Substituting these into $c_i^{1-\sigma} = c_D^{1-\sigma} + \mathbb{1}_{Mi} c_M^{1-\sigma}$ and using $\Delta = V(\phi_M)/V(\phi_D)$ yields (3).

The free entry condition is expressed as

$$\int_{\varphi_D}^{\infty} \pi_{Di} dG(\varphi) + \int_{\varphi_X}^{\infty} \pi_{Xi} dG(\varphi) = f_E,$$

Noting that $\pi_{Di} = Bc_i^{1-\sigma}\varphi^{\sigma-1} - f_{Di}$ where $c_i^{1-\sigma}$ satisfies (3), the expected domestic profits are given by

$$\int_{\varphi_D}^{\infty} \pi_{Di} dG(\varphi) = \int_{\varphi_D}^{\varphi_{DM}} \left(Bc_D^{1-\sigma}\varphi^{\sigma-1} - f_D\right) dG(\varphi) + \int_{\varphi_{DM}}^{\infty} \left(B(1+\tau_M^{1-\sigma}\Delta)c_D^{1-\sigma}\varphi^{\sigma-1} - f_D - f_{DM}\right) dG(\varphi).$$

Rearranging and substituting φ_D and φ_{DM} in (4) into the above equality,

$$\int_{\varphi_D}^{\infty} \pi_{Di} dG(\varphi) = f_D \int_{\varphi_D}^{\infty} \left[\left(\frac{\varphi}{\varphi_D}\right)^{\sigma-1} - 1\right] dG(\varphi) + f_{DM} \int_{\varphi_{DM}}^{\infty} \left[\left(\frac{\varphi}{\varphi_{DM}}\right)^{\sigma-1} - 1\right] dG(\varphi).$$

Similarly, the expected export profits are given by

$$\int_{\varphi_X}^{\infty} \pi_{Xi} dG(\varphi) = f_X \int_{\varphi_X}^{\infty} \left[\left(\frac{\varphi}{\varphi_X}\right)^{\sigma-1} - 1\right] dG(\varphi) + f_{XM} \int_{\varphi_{XM}}^{\infty} \left[\left(\frac{\varphi}{\varphi_{XM}}\right)^{\sigma-1} - 1\right] dG(\varphi).$$

Using the definition of $J(\varphi_c)$ yields (7).

To derive the domestic output share in (8), note that λ_X is given

$$\lambda_X = \frac{\int_i r_{Di} di}{\int_i r_{Di} di + \int_i r_{Xi} di} = \frac{R_D}{R_D + R_X}.$$

The aggregate revenue in the denominator is expressed as

$$R = M_E \int_{\varphi_D}^{\infty} r_{Di} dG(\varphi) + M_E \int_{\varphi_X}^{\infty} r_{Xi} dG(\varphi).$$

Noting that $r_{Di} = \sigma B c_i^{1-\sigma} \varphi^{\sigma-1}$where $c_i^{1-\sigma}$ satisfies (3), the aggregate domestic revenue is given by

$$R_D = M_E \int_{\varphi_D}^{\varphi_{DM}} \left(\sigma B c_D^{1-\sigma} \varphi^{\sigma-1}\right) dG(\varphi) + M_E \int_{\varphi_{DM}}^{\infty} \left(\sigma B (1 + \tau_M^{1-\sigma} \Delta) c_D^{1-\sigma} \varphi^{\sigma-1}\right) dG(\varphi).$$

Using the definition of $V(\varphi_c)$ and rearranging, this aggregate revenue is rewritten as

$$R_D = M_E \sigma B c_D^{1-\sigma} (V(\varphi_D) + \tau_M^{1-\sigma} \Delta V(\varphi_{DM})). \tag{B.1}$$

Similarly, the aggregate export revenue is

$$R_X = M_E \sigma B (\tau_X c_D)^{1-\sigma} (V(\varphi_X) + \tau_M^{1-\sigma} \Delta V(\varphi_{XM})). \tag{B.2}$$

Substituting these in $\lambda_X = R_D/R$ yields (8).

To derive the domestic input share in (9), note that λ_M is given by

$$\lambda_M = \frac{\int_i e_{Di} di}{\int_i e_{Di} di + \int_i e_{Mi} di} = \frac{E_D}{E_D + E_M},$$

where $e_{Di} = \int_v e_{Div} dv$ and $e_{Mi} = \int_v e_{Miv} dv$. The aggregate expenditure in the denominator is expressed as

$$E = M_E \int_{\varphi_D}^{\infty} e_{Di} dG(\varphi) + M_E \int_{\varphi_{DM}}^{\infty} e_{Mi} dG(\varphi).$$

Note that $e_{Div} = p_{Dv}^{1-\sigma} c_i^{\sigma-1} e_i$ and $e_{Miv} = p_{Mv}^{1-\sigma} c_i^{\sigma-1} e_i$ where $e_i = c_i q_i / \varphi$ is written as

$$e_i = \frac{c_i}{\varphi} (q_{Di} + \mathbb{1}_{Xi} \tau_X q_{Xi}),$$

where $\mathbb{1}_{Xi}$ is an indicator function which takes the value of one if firm i exports final goods and zero otherwise. Substituting consumers' demand q_{Di}, q_{Xi} into e_i and subsequently using this in e_{Div}, e_{Miv},

$$\begin{aligned} e_{Div} &= (\sigma - 1) B p_{Dv}^{1-\sigma} \varphi^{\sigma-1} (1 + \mathbb{1}_{Xi} \tau_X^{1-\sigma}), \\ e_{Miv} &= (\sigma - 1) B p_{Mv}^{1-\sigma} \varphi^{\sigma-1} (1 + \mathbb{1}_{Xi} \tau_X^{1-\sigma}). \end{aligned} \tag{B.3}$$

Recalling that $\int_v p_{Dv}^{1-\sigma} dv = c_D^{1-\sigma}$ and $\int_v p_{Mv}^{1-\sigma} dv = c_M^{1-\sigma}$, aggregation of e_{Div}, e_{Miv} yileds

$$\begin{aligned} e_{Di} &= (\sigma - 1) B c_D^{1-\sigma} \varphi^{\sigma-1} (1 + \mathbb{1}_{Xi} \tau_X^{1-\sigma}), \\ e_{Mi} &= (\sigma - 1) B c_M^{1-\sigma} \varphi^{\sigma-1} (1 + \mathbb{1}_{Xi} \tau_X^{1-\sigma}). \end{aligned}$$

Then, the aggregate expenditure of domestic inputs is expressed as

$$E_D = M_E(\sigma - 1)Bc_D^{1-\sigma}(V(\varphi_D) + \tau_X^{1-\sigma}V(\varphi_X)). \tag{B.4}$$

Similarly, the aggregate expenditure of foreign inputs is expressed as

$$E_M = M_E(\sigma - 1)Bc_M^{1-\sigma}(V(\varphi_{DM}) + \tau_X^{1-\sigma}V(\varphi_{XM})). \tag{B.5}$$

Substituting these and (3) into $\lambda_M = E_D/E$ yields (9).

B.1.2 Suppliers

Using (B.3), (B.4) and (B.5), supplier v's revenues $r_{Dv} = \int_i e_{Div}di$ and $r_{Mv} = \int_i e_{Miv}di$ are given by

$$r_{Dv} = p_{Dv}^{1-\sigma}c_D^{\sigma-1}E_D,$$
$$r_{Mv} = p_{Mv}^{1-\sigma}c_M^{\sigma-1}E_M.$$

Moreover, it follows from (B.4) and (B.5) that

$$E_M = \left(\frac{c_M}{c_D}\right)^{1-\sigma}\Lambda_M E_D,$$

and hence

$$r_{Mv} = p_{Mv}^{1-\sigma}\Lambda_M c_D^{\sigma-1}E_D.$$

Using $p_{Dv} = \frac{\sigma}{\sigma-1}\frac{1}{\phi}$ and $p_{Mv} = \frac{\sigma}{\sigma-1}\frac{\tau_M}{\phi}$ and the definition of A yields domestic and export revenues r_{Dv}, r_{Mv}. The zero profit cutoff condition (10) follows immediately from that expression.

The free entry condition is expressed as

$$\int_{\phi_D}^{\infty}\pi_{Dv}dG(\phi) + \int_{\phi_M}^{\infty}\pi_{Mv}dG(\phi) = k_E.$$

Noting that $\pi_{Dv} = A\phi^{\sigma-1} - k_D$, the expected domestic profits are given by

$$\int_{\phi_D}^{\infty}\pi_{Dv}dG(\phi) = \int_{\phi_D}^{\infty}\left(A\phi^{\sigma-1} - k_D\right)dG(\phi).$$

Substituting ϕ_D in (10) into the above equality,

$$\int_{\phi_D}^{\infty}\pi_{Dv}dG(\phi) = k_D\int_{\phi_D}^{\infty}\left[\left(\frac{\phi}{\phi_D}\right)^{\sigma-1} - 1\right]dG(\phi).$$

Similarly, the expected export profits are given by

$$\int_{\phi_M}^{\infty}\pi_{Mv}dG(\phi) = k_M\int_{\phi_M}^{\infty}\left[\left(\frac{\phi}{\phi_M}\right)^{\sigma-1} - 1\right]dG(\phi).$$

Using the definition of $J(\phi_c)$ yields (12).

To derive the domestic input share in (9), note that λ_M is alternatively given by

$$\lambda_M = \frac{\int_v r_{Dv} dv}{\int_v r_{Dv} dv + \int_v r_{Mv} dv} = \frac{E_D}{E_D + E_M}.$$

Noting that $r_{Dv} = \sigma A \phi^{\sigma-1}$, the aggregate domestic revenue is given by

$$E_D = N_E \int_{\phi_D}^{\infty} \left(\sigma A \phi^{\sigma-1}\right) dG(\phi).$$

Using the definition of $V(\phi_c)$, this revenue is rewritten as

$$E_D = N_E \sigma A V(\phi_D). \tag{B.6}$$

Similarly, the aggregate export revenue is expressed as

$$E_M = N_E \sigma A \tau_M^{1-\sigma} \Lambda_M V(\phi_M). \tag{B.7}$$

Substituting these in $\lambda_M = E_D/E$ yields (9).

B.1.3 Labor Market and Welfare

To define the labor market clearing condition of the economy, consider the aggregate amount of labor used for entry and production at the downstream stage ($\int_i l_i di = \int_i l_i^e di + \int_i l_i^p di$). This amount is written as

$$\int_i l_i di = M_E f_E + M_E \int_{\varphi_D}^{\infty} f_{Di} dG(\varphi) + M_E \int_{\varphi_X}^{\infty} f_{Xi} dG(\varphi).$$

Every entrant incurs fixed entry costs f_E, and firm i above φ_D (φ_X) also incurs fixed overhead (export) costs f_{Di} (f_{Xi}) to serve the domestic (foreign) market. Note that production worker is used only for fixed costs at the downstream stage since we have assumed that labor is not hired for transforming inputs into final goods. From $\pi_{Di} = \frac{r_{Di}}{\sigma} - f_{Di}$, $\pi_{Xi} = \frac{r_{Xi}}{\sigma} - f_{Xi}$ and the free entry condition (7), we get

$$\int_i l_i di = R - E. \tag{B.8}$$

Similarly the aggregate amount of labor at the upstream stage ($\int_v l_v dv = \int_v l_v^e dv + \int_v l_v^p dv$) is written as

$$\int_v l_v dv = N_E k_E + N_E \int_{\phi_D}^{\infty} l_{Dv}^p G(\phi) + N_E \int_{\phi_M}^{\infty} l_{Mv}^p dG(\phi),$$

where $l_{Dv}^p = k_D + \frac{x_{Dv}}{\phi}$ and $l_{Mv}^p = k_M + \frac{\tau_M x_{Mv}}{\phi}$. Every entrant incurs fixed entry costs k_E, and supplier v above ϕ_D (ϕ_M) also hire labor used for production l_{Dv}^p (l_{Mv}^p) to serve the domestic (foreign) market. This implies that production worker is used for both variable costs and fixed costs at the upstream stage. From $\pi_{Dv} = \frac{r_{Dv}}{\sigma} - k_D$, $\pi_{Mv} = \frac{r_{Mv}}{\sigma} - k_M$ and the free entry condition (12), we get

$$\int_v l_v dv = E. \tag{B.9}$$

Substituting (B.8) and (B.9) into $\int_i l_i di + \int_v l_v dv = L$, we get the standard labor market condition $R = L$. Moreover, noting $R = R_D + R_X$ in (B.1) and (B.2) as well as $E = E_D + E_M$ in (B.4) and (B.5),

$$E = \left(\frac{\sigma - 1}{\sigma}\right) R. \tag{B.10}$$

Substituting this into (B.8) and (B.9) yields the aggregate amount of labor allocated to each production stage.

To get the mass of entrants at each production stage, consider the mass of entrants at the downstream stage. Substituting φ_c in (4) into (B.1) and (B.2), $R = R_D + R_X$ is expressed as

$$R = M_E \sigma \sum_c f_c \varphi_c^{1-\sigma} V(\varphi_c).$$

Moreover, using $R = L$ and rewriting this gives us the mass of entrants at the downstream stage:

$$M_E = \frac{L}{\sigma \sum_c f_c \varphi_c^{1-\sigma} V(\varphi_c)}. \tag{B.11}$$

Similarly, substituting ϕ_c in (10) into (B.6) and (B.7), $E = E_D + E_M$ is expressed as

$$E = N_E \sigma \sum_c k_c \phi_c^{1-\sigma} V(\phi_c).$$

Moreover, using (B.10) and rewriting this gives us the mass of entrants at the upstream stage:

$$N_E = \left(\frac{\sigma - 1}{\sigma}\right) \frac{L}{\sigma \sum_c k_c \phi_c^{1-\sigma} V(\phi_c)}. \tag{B.12}$$

Regarding welfare per worker defined as $W \equiv U/L$, it follows from $Q \equiv U$ and $PQ = R = L$ that

$$W = \frac{Q}{L} = \frac{1}{P}.$$

Noting that $w = 1$, welfare per worker is equivalent to the real wage. To get P, solving the first equality in (4) for B, substituting it into the definition of the final goods market demand B and using $R = L$,

$$\frac{1}{P} = \left(\frac{\sigma - 1}{\sigma}\right) \left(\frac{L}{\sigma f_D}\right)^{\frac{1}{\sigma-1}} \frac{\varphi_D}{c_D}, \tag{B.13}$$

which depends on both firms' domestic productivity cutoff φ_D and firms' unit costs c_D. To get c_D, solving the first equality in (10) for A, substituting it into the definition of the input market demand A, and using (B.10),

$$\frac{1}{c_D} = \left(\frac{\sigma - 1}{\sigma}\right)^{\frac{\sigma}{\sigma-1}} \left(\frac{L}{\sigma k_D}\right)^{\frac{1}{\sigma-1}} \phi_D \lambda_M^{\frac{1}{\sigma-1}}. \tag{B.14}$$

Combining (B.13) and (B.14), we can express welfare as

$$W = \left(\frac{\sigma - 1}{\sigma}\right)^{\frac{2\sigma-1}{\sigma-1}} \left(\frac{L}{\sigma f_D}\right)^{\frac{1}{\sigma-1}} \left(\frac{L}{\sigma k_D}\right)^{\frac{1}{\sigma-1}} \varphi_D \phi_D \lambda_M^{\frac{1}{\sigma-1}}.$$

Rewriting this gives us the welfare expression in (13).

Though we have derived welfare per worker from the definitions of the market demands A, B above, it is possible to derive it from the price index. From the output prices p_{Di}, p_{Xi}, the price index is given by

$$P^{1-\sigma} = M_E \int_{\varphi_D}^{\infty} \left(\frac{\sigma}{\sigma-1} \frac{c_i}{\varphi} \right)^{1-\sigma} dG(\varphi) + M_E \int_{\varphi_X}^{\infty} \left(\frac{\sigma}{\sigma-1} \frac{\tau_X c_i}{\varphi} \right)^{1-\sigma} dG(\varphi).$$

Substituting $c_i^{1-\sigma}$ from (3) and aggregating over the relevant productivity ranges, this is expressed as

$$P^{1-\sigma} = M_E \left(\frac{\sigma c_D}{\sigma-1} \right)^{1-\sigma} \left[V(\varphi_D) + \tau_M^{1-\sigma} \Delta V(\varphi_{DM}) + \tau_X^{1-\sigma} V(\varphi_X) + (\tau_X \tau_M)^{1-\sigma} \Delta V(\varphi_{XM}) \right].$$

Further, using φ_c in (4) and rearranging, the price index is simply rewritten as

$$P^{1-\sigma} = \frac{M_E}{B} \left(\frac{\sigma}{\sigma-1} \right)^{1-\sigma} \sum_c f_c \varphi_c^{1-\sigma} V(\varphi_c).$$

Finally, substituting M_E from (B.11) and $B = f_D(\varphi_D/c_D)^{1-\sigma}$ from (4) gives us the same expression as (B.13). On the other hand, the unit costs of non-importing firms are given by

$$c_D^{1-\sigma} = N_E \left(\frac{\sigma}{\sigma-1} \right)^{1-\sigma} V(\phi_D).$$

Substituting N_E from (B.12) and rearranging, this is expressed as

$$c_D^{1-\sigma} = \left(\frac{\sigma}{\sigma-1} \right)^{-\sigma} \frac{k_D \phi_D^{1-\sigma} V(\varphi_D)}{\sum_c k_c \phi_c^{1-\sigma} V(\phi_c)} \frac{L}{\sigma k_D \phi_D^{1-\sigma}}.$$

Finally, substituting $\frac{k_D \phi_D^{1-\sigma} V(\varphi_D)}{\sum_c k_c \phi_c^{1-\sigma} V(\phi_c)} = \frac{\beta}{\beta+1} = \lambda_M$ from Appendix A.1 gives us the same expression as (B.14).

B.1.4 Trade Flows

Under the Pareto distribution, $J(\varphi_c) = \int_{\varphi_c}^{\infty} \left[\left(\frac{\varphi}{\varphi_c} \right)^{\sigma-1} - 1 \right] dG(\varphi)$ and $V(\varphi_c) = \int_{\varphi_c}^{\infty} \varphi^{\sigma-1} dG(\varphi)$ are simple power functions of the productivity cutoff φ_c:

$$\begin{aligned} J(\varphi_c) &= \frac{\sigma-1}{\theta} \frac{1}{\varphi_c^{\sigma-1+\theta}}, \\ V(\varphi_c) &= \frac{\sigma-1+\theta}{\theta} \frac{1}{\varphi_c^{\theta}}. \end{aligned} \tag{B.15}$$

which satisfy $J(\varphi_c) = \frac{\sigma-1}{\sigma-1+\theta} \varphi_c^{1-\sigma} V(\varphi_c)$.

To compute the mass of entrants M_E, N_E under the Pareto distribution, using (B.15) in the mass of entrants at the downstream stage (B.11),

$$M_E = \frac{L}{\frac{\sigma(\sigma-1+\theta)}{\theta} \left(\sum_c f_c \varphi_c^{-(\sigma-1+\theta)} \right)}.$$

Further, applying the Pareto distribution to the free entry condition (7),

$$\frac{\sigma-1}{\theta} \left[\sum_c f_c \varphi_c^{-(\sigma-1+\theta)} \right] = f_E.$$

Combining the two expressions, we have that the mass of entrants is proportional to labor endowment L:

$$M_E = \frac{\sigma - 1}{\sigma(\sigma - 1 + \theta)} \frac{L}{f_E}. \tag{B.16}$$

Similarly, using (B.15) in the mass of entrants at the upstream stage (B.12) and applying the Pareto distribution to the free entry condition (12), we also have

$$N_E = \frac{(\sigma - 1)^2}{\sigma^2(\sigma - 1 + \theta)} \frac{L}{k_E}. \tag{B.17}$$

To derive the gravity equation (21), let input trade flows $E_M = \int_v r_{Mv} dv$ decompose into the average sales per supplier and the mass of suppliers:

$$\begin{aligned} E_M &= \frac{1}{1 - G(\phi_M)} \int_{\phi_M}^{\infty} r_{Mv} dG(\phi_M) \;\times\; [1 - G(\phi_M)] N_E \\ &= \frac{\sigma(\sigma - 1 + \theta)}{\theta} k_M \;\times\; \left(\frac{1}{\phi_M}\right)^{\sigma - 1 + \theta} N_E, \end{aligned}$$

where the second equality follows from using (10) and (B.15) in suppliers' export revenues $r_{Mv} = \sigma A \tau_M^{1-\sigma} \Lambda_M \phi^{\sigma - 1}$. This decomposition in turn can be rearranged as

$$\begin{aligned} E_M &= \frac{\sigma(\sigma - 1 + \theta)}{\theta} k_M \;\times\; \phi_M^{-(\sigma - 1 + \theta)} \left(\frac{(\sigma - 1)^2}{\sigma^2(\sigma - 1 + \theta)} \frac{L}{k_E} \right) && \text{(using (B.17))} \\ &= \left(\frac{(\sigma - 1)^2}{\sigma \theta k_E} \right) L k_M \left(\frac{k_M}{A \tau_M^{1-\sigma} \Lambda_M} \right)^{-\frac{\sigma - 1 + \theta}{\sigma - 1}} && \text{(using (10))} \\ &= \psi_M \Lambda_M^{\frac{\sigma - 1 + \theta}{\sigma - 1}} L A^{\frac{\sigma - 1 + \theta}{\sigma - 1}} \tau_M^{-(\sigma - 1 + \theta)} k_M^{-\frac{\theta}{\sigma - 1}}. \end{aligned}$$

Similarly, output trade flows are decomposed into

$$\begin{aligned} R_X &= \frac{1}{\eta_X} \frac{\sigma(\sigma - 1 + \theta)}{\theta} f_X \;\times\; (\varphi_X)^{-(\sigma - 1 + \theta)} \left(\frac{\sigma - 1}{\sigma(\sigma - 1 + \theta)} \frac{L}{f_E} \right) && \text{(using (B.16))} \\ &= \frac{1}{\eta_X} \left(\frac{\sigma - 1}{\theta f_E} \right) L f_X \left(\frac{f_X}{B(\tau_X c_D)^{1-\sigma}} \right)^{-\frac{\sigma - 1 + \theta}{\sigma - 1}} && \text{(using (4))} \\ &= \frac{\psi_X}{\eta_X} L B^{\frac{\sigma - 1 + \theta}{\sigma - 1}} (\tau_X c_D)^{-(\sigma - 1 + \theta)} f_X^{-\frac{\theta}{\sigma - 1}}. \end{aligned}$$

This gives us the gravity equation expressions in (21).

To get the output trade elasticity from the gravity equation, dividing aggregate foreign expenditure E_M in (B.5) by aggregate domestic expenditure E_D in (B.4) and subsequently applying (B.15),

$$\frac{E_M}{E_D} = \tau_M^{-(\sigma - 1 + \theta)} \Lambda_M^{\frac{\sigma - 1 + \theta}{\sigma - 1}} \left(\frac{k_M}{k_D} \right)^{-\frac{\theta}{\sigma - 1}}.$$

Taking the log and differentiating E_M / E_D with respect to τ_M,

$$\varepsilon_M = \sigma - 1 + \theta + \left(\frac{\sigma - 1 + \theta}{\sigma - 1} \frac{d \ln \Lambda_M}{d \ln \tau_M} \right).$$

Moreover, noting that $\Lambda_M = \frac{\mu_D V(\varphi_{DM})}{\mu_M V(\varphi_D)}$ and solving for Δ and Λ_M,

$$\begin{aligned}
\Delta^{\frac{\sigma-1+\theta}{\sigma-1}} &= \left[\tau_M^{-(\sigma-1+\theta)} \frac{\mu_D}{\mu_M} \left(\frac{f_{DM}}{f_D}\right)^{-\frac{\theta}{\sigma-1}} \left(\frac{k_M}{k_D}\right)^{-1}\right]^{\frac{\theta}{\sigma-1-\theta}}, \\
\Lambda_M^{\frac{\sigma-1+\theta}{\sigma-1}} &= \left[\tau_M^{-(\sigma-1+\theta)} \left(\frac{\mu_D}{\mu_M}\right)^{\frac{\sigma-1}{\theta}} \left(\frac{f_{DM}}{f_D}\right)^{-1} \left(\frac{k_M}{k_D}\right)^{-\frac{\theta}{\sigma-1}}\right]^{\frac{\theta}{\sigma-1-\theta}}.
\end{aligned} \tag{B.18}$$

While Λ_M in (B.18) depends not only on τ_M directly but also on μ_D/μ_M indirectly, the definition of μ_D, μ_M implies that $\frac{d\ln(\mu_D/\mu_M)}{d\ln\tau_M} = 0$ and only the direct changes of τ_M remains. This yields the same expression of the input trade elasticity in (19). Similarly, dividing R_X in (B.2) by R_D in (B.1) and subsequently applying (B.15),

$$\frac{R_X}{R_D} = \tau_X^{-(\sigma-1+\theta)} \frac{\eta_D}{\eta_X} \left(\frac{f_X}{f_D}\right)^{-\frac{\theta}{\sigma-1}}.$$

Taking the log and differentiating R_X/R_D with respect to τ_X,

$$\varepsilon_X = \sigma - 1 + \theta - \left(\frac{d\ln(\eta_D/\eta_X)}{d\ln\tau_X}\right).$$

Moreover, from the definition of η_D, η_X, it follows that

$$-\left(\frac{d\ln(\eta_D/\eta_X)}{d\ln\tau_X}\right) = \theta\left(\frac{\sigma-1+\theta}{\sigma-1-\theta}\right)(\eta_D - \eta_X)(\mu_D - \mu_M).$$

This yields the same expression of the output trade elasticity in (20).

To show the trade elasticities capturing the indirect effect, note that

$$\begin{aligned}
\tilde{\varepsilon}_M &= -\frac{d\ln(E_M/E_D)}{d\ln\tau_X} = -\frac{d\ln\Delta}{d\ln\tau_X} - \frac{d\ln\Lambda_M}{d\ln\tau_X}, \\
\tilde{\varepsilon}_X &= -\frac{d\ln(R_X/R_D)}{d\ln\tau_M} = -\frac{d\ln\Lambda_X}{d\ln\tau_M},
\end{aligned}$$

which shows that there is only the extensive margin elasticity. While the closed-form solutions of Δ and Λ_M are given in (B.18), applying (B.15) to $\Lambda_X = \frac{\eta_D V(\varphi_X)}{\eta_X V(\varphi_D)}$, we get

$$\Lambda_X = \frac{\eta_D}{\eta_X}\left(\frac{\tau_X^{\sigma-1} f_X}{f_D}\right)^{-\frac{\theta}{\sigma-1}}.$$

It can be easily shown that

$$\begin{aligned}
-\frac{d\ln\Delta}{d\ln\tau_X} &= \left(\frac{\theta(\sigma-1)}{\sigma-1-\theta}\right)(\mu_D - \mu_M), \\
-\frac{d\ln\Lambda_M}{d\ln\tau_X} &= \left(\frac{(\sigma-1)^2}{\sigma-1-\theta}\right)(\mu_D - \mu_M), \\
-\frac{d\ln\Lambda_X}{d\ln\tau_M} &= \left(\frac{(\sigma-1)(\sigma-1+\theta)}{\sigma-1-\theta}\right)(\eta_D - \eta_X).
\end{aligned}$$

This gives us the expressions of the trade elasticities in (22).

B.2 Condition (18)

This section provides detailed explanations for (18). We show that (18) is related to complementarity between exporting and importing in revenues in general, and in fixed costs in particular under the Pareto distribution.

To show this, we find it useful to introduce a firm revenue share which is directly observable from the data. Let r_{Di}/R denote firm i's domestic revenue share in the domestic market. From Section 2.2, firm i's domestic revenues are $r_{Di} = \sigma B c_i^{1-\sigma}\varphi^{\sigma-1}$ where firm i's unit costs are (3). Moreover, from Section 2.4, the aggregate revenue of firms equals the fixed labor endowment, $R = L$. Then, let s_D (s_{DM}) denote all non-importing (all importing) firms' revenue share in the domestic market. From the sorting patterns in the domestic market, s_D (s_{DM}) is obtained by aggregating r_{Di}/R over the productivity range between φ_D and φ_{DM} (above φ_{DM}) among a mass M_E of entrants: $s_D = M_E \int_{\varphi_D}^{\varphi_{DM}} B\sigma\varphi^{\sigma-1}dG(\varphi)/R$ and $s_{DM} = M_E \int_{\varphi_{DM}}^{\infty} B\sigma(1+\tau_M^{1-\sigma}\Delta)\varphi^{\sigma-1}dG(\varphi)/R$. Similarly, let r_{Xi}/R denote firm i's export revenue share in the foreign market, while let s_X (s_{XM}) denote all non-importing (all importing) firms' revenue share in the foreign market. Since firm i's export revenues are $r_{Xi} = \sigma B(\tau_X c_i)^{1-\sigma}\varphi^{\sigma-1}$ and $R = L$ in the foreign market, we can calculate revenue shares s_X and s_{XM} from the souring patterns in the foreign market. Using $V(\varphi_c)$, these revenue shares are respectively given by

$$
\begin{aligned}
s_D &= \frac{M_E\sigma B c_D^{1-\sigma}[V(\varphi_D) - V(\varphi_{DM})]}{R}, \\
s_{DM} &= \frac{M_E\sigma B c_D^{1-\sigma}(1+\tau_M^{1-\sigma}\Delta)V(\varphi_{DM})}{R}, \\
s_X &= \frac{M_E\sigma B(\tau_X c_D)^{1-\sigma}[V(\varphi_X) - V(\varphi_{XM})]}{R}, \\
s_{XM} &= \frac{M_E\sigma B(\tau_X c_D)^{1-\sigma}(1+\tau_M^{1-\sigma}\Delta)V(\varphi_{XM})}{R}.
\end{aligned}
$$

Note that s_X and s_{XM} also represent export revenue share in the domestic market under country symmetry, and hence the revenue shares must satisfy $\sum_c s_c = 1$ (from $R = \int_i r_{Di}di + \int_i r_{Xi}di = L$). Taking the ratio of revenue share between importing firms and non-importing firms by classifying their export status,

$$
\frac{s_{XM}}{s_X} > \frac{s_{DM}}{s_D} \quad\Longleftrightarrow\quad \frac{V(\varphi_{XM})}{V(\varphi_X)} > \frac{V(\varphi_{DM})}{V(\varphi_D)},
$$

where the latter inequality is the same as (18).

Next consider the relationship between $V(\varphi_c)/V(\varphi_{c'})$ and $s_c/s_{c'}$. On the one hand, $V(\varphi_{DM})/V(\varphi_D)$ is the relative revenues of importing firms among non-exporting firms in the domestic market, while $V(\varphi_{XM})/V(\varphi_X)$ is the relative revenues of importing firms among exporting firms in the same market. On the other hand, s_{DM}/s_D is the relative revenue share of importing firms among non-exporting firms in the domestic market, while s_{XM}/s_X is the relative revenue share of importing firms among exporting firms in the same market. The interpretation of these variables is slightly different, but both of them clearly capture the relative revenues in the domestic market. Thus (18) embodies the idea that exporting and importing exhibit complementarity that happens via an increase in firm revenues: firms engaging in both exporting and importing have relatively larger revenues (or revenue share) than firms engaging in exporting only or importing only.

Finally, note that neither $V(\varphi_c)/V(\varphi_{c'})$ nor $s_c/s_{c'}$ directly depends on fixed costs, which indirectly affect these variables through productivity cutoffs among others. When we posit the Pareto distribution, the relative revenues are expressed as $V(\varphi_c)/V(\varphi_{c'}) = (\varphi_{c'}/\varphi_c)^{\theta}$, which are proportional to the relative fixed costs $f_{c'}/f_c$ from (4). In this special case, complementarity happens via a reduction in fixed costs.

B.3 Extensions

This section presents detailed results when a Pareto share parameter differs across the two production stages. Clear analytical results are provided for both Propositions 1 and 2 in this general case.

Proposition 1. We first consider the impact of input trade liberalization on the domestic productivity cutoffs. Differentiating and solving the equilibrium conditions simultaneously, reductions in input trade costs have the following impact on the domestic productivity cutoffs (its proof is very similar to that in Appendix A.1):

$$\begin{aligned} d\ln\phi_D &= -\left(\frac{(\sigma-1)(\sigma-1+\vartheta)(1-\lambda_M)}{(\sigma-1)^2-\theta\vartheta}\right) d\ln\tau_M, \\ d\ln\varphi_D &= -\left(\frac{(\sigma-1)(\sigma-1+\theta)(1-\lambda_M)}{(\sigma-1)^2-\theta\vartheta}\right) d\ln\tau_M, \end{aligned} \tag{B.19}$$

where we now assume $(\sigma-1)^2-\theta\vartheta>0$ (instead of (16)). It is not surprising to see that if the Pareto shape parameter is common between the two production stages so that $\theta=\vartheta$, (B.19) collapses to (15) in the baseline model. In contrast to (15) where changes at the domestic productivity cutoffs are exactly the same between the upstream and downstream stages, changes in these two cutoffs are different in this general case. In particular, it follows from (B.19) that the changes in φ_D is greater than those in ϕ_D if and only if $\theta>\vartheta$, which means that when the upstream stage is more homogeneous than the downstream stage, input trade liberalization has a greater impact on firms than on suppliers. Intuitively, when $\theta>\vartheta$, suppliers are more sensitive to changes in trade costs than firms and reductions in τ_M (by input trade liberalization) allow suppliers to ship their input relatively more easily than $\theta=\vartheta$; since more inputs are available at the downstream stage, competition among firms is tougher (via lower production costs from an expansion of input varieties), and an increase is greater for φ_D than for ϕ_D. Nonetheless, (B.19) shows that input trade liberalization increases both cutoffs.

Let us next investigate the impact of output trade liberalization on the domestic productivity cutoffs. It can be shown that reductions in output trade costs have the following impact on the domestic productivity cutoffs (its proof is very similar to that in Appendix A.2):

$$\begin{aligned} d\ln\phi_D &= -\left(\frac{(\sigma-1)(\sigma-1+\vartheta)(\mu_D-\mu_M)(1-\lambda_M)}{(\sigma-1)^2-\theta\vartheta}\right) d\ln\tau_X, \\ d\ln\varphi_D &= -\left(1-\lambda_X+\frac{\theta(\sigma-1+\vartheta)(\mu_D-\mu_M)(1-\lambda_M)}{(\sigma-1)^2-\theta\vartheta}\right) d\ln\tau_X. \end{aligned} \tag{B.20}$$

As in (B.19), if the shape parameter is common so that $\theta=\vartheta$, (B.20) collapses to (17) in the baseline model. Note, however, that the changes in the cutoffs are different in both the baseline case and the general case. Thus, if $(\sigma-1)^2-\theta\vartheta>0$ and (18) hold, (B.20) shows that input trade liberalization increases both cutoffs.

Comparison of the impact on the domestic productivity cutoff in the *upstream* stage reveals that

$$\left|\frac{d\ln\phi_D}{d\ln\tau_M}\right| > \left|\frac{d\ln\phi_D}{d\ln\tau_X}\right|,$$

and input trade liberalization always generates the greater productivity gains than output trade liberalization. On the other hand, changes in the domestic productivity cutoff in the *downstream* stage satisfy

$$\left|\frac{d\ln\varphi_D}{d\ln\tau_M}\right| > \left|\frac{d\ln\varphi_D}{d\ln\tau_X}\right| \quad \Longleftrightarrow \quad ((\sigma-1)^2-\theta\vartheta)(\lambda_X-\lambda_M)+\theta(\sigma-1+\vartheta)(1-\mu_D+\mu_M)(1-\lambda_M)>0,$$

which shows that the sufficient condition for this inequality is $\lambda_X \geq \lambda_M$, as in the baseline model. Therefore, while the impacts of trade liberalization on the cutoffs require some qualification (i.e., the changes in the cutoffs to input trade costs are no longer the same between the stages), the results in Lemmas 1 and 2 continue to hold even with different shape parameters.

Turning to the impact on trade flows, consider the impact of input trade liberalization on input trade flows. If the degree of heterogeneity in productivity is different by stage, the exporter and importer extensive margin elasticities are also different by stage, which comes from the difference in (B.19):

$$
\begin{aligned}
-\frac{d\ln\Delta}{d\ln\tau_M} &= \frac{\theta(\sigma-1)(\sigma-1+\vartheta)}{(\sigma-1)^2-\theta\vartheta}, \\
-\frac{d\ln\Lambda_M}{d\ln\tau_M} &= \frac{\vartheta(\sigma-1)(\sigma-1+\theta)}{(\sigma-1)^2-\theta\vartheta}.
\end{aligned}
$$

It can be easily confirmed that the two extensive margin elasticities are increasing in θ, which means that as the upstream stage is more homogeneous than the downstream stage, the extensive margin elasticity increases not only at the upstream stage but also at the downstream stage. The result is intuitive: as θ is larger, suppliers are more sensitive to changes in trade costs and reductions in τ_M induce more suppliers to start exporting at the upstream stage; at the same time, this increase in input trade improves the range of input available to firms, inducing more firms to start importing at the downstream stage. Summing up the intensive margin elasticity $\sigma - 1$ and the above extensive margin elasticities, the input trade elasticity is given by

$$
\varepsilon_M = \frac{(\sigma-1)(\sigma-1+\theta)(\sigma-1+\vartheta)}{(\sigma-1)^2-\theta\vartheta}, \tag{B.21}
$$

which is also increasing in θ and thus a more homogeneous upstream stage (relative to the baseline model with $\theta = \vartheta$) leads to an increase in the input trade elasticity.

Next consider the impact of output trade liberalization on output trade flows. As for the exporter extensive margin at the downstream stage, the market share of exporting firms Λ_X depends on the supplier distribution $G(\phi)$ through $\Delta = V(\phi_M)/V(\phi_D)$ as well as the firm distribution $G(\varphi)$ through $V(\varphi_c)$. Since these distribution functions are now different due to the different shape parameter by stage, the exporter extensive margin at the downstream stage has the two extensive margin elasticities:

$$
-\frac{d\ln\Lambda_X}{d\ln\tau_X} = \vartheta\left[1+\left(\frac{\theta(\sigma-1+\vartheta)}{(\sigma-1)^2-\theta\vartheta}\right)(\eta_D-\eta_X)(\mu_D-\mu_M)\right] + \theta\left(\frac{(\sigma-1)(\sigma-1+\vartheta)}{(\sigma-1)^2-\theta\vartheta}\right)(\eta_D-\eta_X)(\mu_D-\mu_M).
$$

Thus, reductions in τ_X allow not only more firms to start exporting at the downstream stage directly (first term), but also more suppliers to start exporting at the upstream stage indirectly (second term). As a consequence, the output trade elasticity in (20) is now given by

$$
\varepsilon_X = \sigma - 1 + \vartheta + \theta\left(\frac{(\sigma-1+\vartheta)^2}{(\sigma-1)^2-\theta\vartheta}\right)(\eta_D-\eta_X)(\mu_D-\mu_M). \tag{B.22}
$$

This expression shows that the output trade elasticity is decomposed into the intensive margin elasticity $\sigma - 1$, the extensive margin elasticity at the downstream stage ϑ and the extensive margin elasticity at the upstream stage θ weighted by some composite term which is positive if and only if (18) holds. As in the input trade elasticity ε_M, (B.22) reveals that the output trade elasticity ε_X is also increasing in θ and hence a more homogeneous upstream stage leads to an increase not only in input trade flows but also in output trade flows.

The mechanism behind the result is very similar to before: as θ is larger, suppliers are more sensitive to changes in trade costs and reductions in τ_X induce more suppliers to start exporting at the upstream stage; at the same time, this increase in input trade improves the range of inputs available to firms, which induces more firms to start exporting the downstream stage by exploiting the love-of-variety effect. Despite this new outcome in the extensive margin, simple inspection of (B.21) and (B.22) immediately reveals that the trade elasticity results in Proposition 1 continue to hold in this general case with different shape parameters:

$$\varepsilon_M > \varepsilon_X.$$

The gravity equations in (21) are now expressed as

$$\begin{aligned} E_M &= \psi_M \Lambda_M^{\frac{\sigma-1+\vartheta}{\sigma-1}} L A^{\frac{\sigma-1+\theta}{\sigma-1}} \tau_M^{-(\sigma-1+\theta)} k_M^{-\frac{\theta}{\sigma-1}}, \\ R_X &= \frac{\psi_X}{\eta_X} L B^{\frac{\sigma-1+\vartheta}{\sigma-1}} (\tau_X c_D)^{-(\sigma-1+\vartheta)} f_X^{-\frac{\vartheta}{\sigma-1}}, \end{aligned}$$

where ψ_M is the same as before and $\psi_X = \frac{\sigma-1}{\vartheta f_E}$. Moreover, we can show that

$$\Lambda_M^{\frac{\sigma-1+\theta}{\sigma-1}} = \left[\tau_M^{-(\sigma-1+\theta)} \left(\frac{\mu_D}{\mu_M}\right)^{\frac{\sigma-1}{\vartheta}} \left(\frac{f_{DM}}{f_D}\right)^{-1} \left(\frac{k_M}{k_D}\right)^{-\frac{\theta}{\sigma-1}}\right]^{\frac{\vartheta(\sigma-1+\theta)}{(\sigma-1)^2-\theta\vartheta}}.$$

This equation shows that the input trade costs indirectly reduces input trade flows through firms' market share (i.e., through Λ_M) with an elasticity of $\frac{\vartheta(\sigma-1+\theta)^2}{(\sigma-1)^2-\theta\vartheta}$. Using the gravity equation of E_M, R_X in the definition of the trade elasticity, we obtain the closed-form solutions of ε_M and ε_X that are exactly the same expressions in (B.21) and (B.22) respectively. This shows that the reduced-form expression (1) continues to hold.

Finally, the trade elasticities capturing the indirect effect in (22) are now given by

$$\begin{aligned} \tilde{\varepsilon}_M &= (\sigma-1)\left(\frac{(\sigma-1+\theta)(\sigma-1+\vartheta)}{(\sigma-1)^2-\theta\vartheta}\right)(\mu_D-\mu_M), \\ \tilde{\varepsilon}_X &= (\sigma-1)\left(\frac{(\sigma-1+\theta)(\sigma-1+\vartheta)}{(\sigma-1)^2-\theta\vartheta}\right)(\eta_D-\eta_X), \end{aligned} \tag{B.23}$$

which shows that when exporting and importing exhibit complementarity so that (18) holds, both $\tilde{\varepsilon}_X$ and $\tilde{\varepsilon}_M$ still have positive values. Thus input trade liberalization increases not only input trade flows directly, but also output trade flows indirectly, and vice versa.

Proposition 2. Before exploring the welfare results when a Pareto share parameter differs by stage, it should be first noticed that the welfare changes in (23) hold even in this general case. Only differences are the changes in the productivity cutoffs as well as domestic input share associated with input and output trade liberalization, which now have two separate shape parameters.

Consider first the impact of input trade liberalization on welfare. The changes in the domestic productivity cutoffs are given by (B.19), while the changes in the domestic input share are the same as before, i.e., $d\ln\lambda_M = (1-\lambda_M)\varepsilon_M d\ln\tau_M$, where ε_M is given by (B.21). Substituting these into (23) and summing up the three terms, the changes in welfare associated with input trade liberalization are given by

$$d\ln W = -\Theta_M(1-\lambda_M)d\ln\tau_M, \tag{B.24}$$

where

$$\Theta_M \equiv \left(\frac{(\sigma-1)(\sigma-1+\vartheta)}{(\sigma-1)^2-\theta\vartheta}\right)+\left(\frac{(\sigma-1)(\sigma-1+\theta)}{(\sigma-1)^2-\theta\vartheta}\right)-\left(\frac{(\sigma-1+\theta)(\sigma-1+\vartheta)}{(\sigma-1)^2-\theta\vartheta}\right).$$

Though (B.24) includes both θ and ϑ reflecting different shape parameters, we can easily check that $\Theta_M = 1$ and the changes in welfare associated with input trade liberalization are simply given by (24). This means that the elasticity of welfare with respect to input trade costs is equivalent to the foreign input share $1-\lambda_M$ as in the baseline model.

Consider next the impact of output trade liberalization on welfare. The changes in the domestic productivity cutoffs are given by (B.20), while the changes in the domestic input share are the same as before, i.e., $d\ln\lambda_M = (1-\lambda_M)\tilde{\varepsilon}_M d\ln\tau_X$, where $\tilde{\varepsilon}_M$ is given by (B.23). Substituting these into (23) and summing up the three terms, the changes in welfare associated with output trade liberalization are given by

$$d\ln W = -\Big[1-\lambda_X+\Theta_X(\mu_D-\mu_M)(1-\lambda_M)\Big]d\ln\tau_X, \tag{B.25}$$

where

$$\Theta_X \equiv \left(\frac{(\sigma-1)(\sigma-1+\vartheta)}{(\sigma-1)^2-\theta\vartheta}\right)+\left(\frac{\theta(\sigma-1+\vartheta)}{(\sigma-1)^2-\theta\vartheta}\right)-\left(\frac{(\sigma-1+\theta)(\sigma-1+\vartheta)}{(\sigma-1)^2-\theta\vartheta}\right).$$

While (B.25) also includes both θ and ϑ, we can easily check that $\Theta_X = 0$ and the changes in welfare associated with output trade liberalization are simply given by (25). This means that the elasticity of welfare with respect to output trade costs is equivalent to the foreign output share $1-\lambda_X$ as in the baseline model.

Finally, it follows immediately from (B.24) and (B.25) that the welfare results in Proposition 2 continue to hold in this general case with different shape parameters:

$$\left|\frac{d\ln W}{d\ln\tau_M}\right| > \left|\frac{d\ln W}{d\ln\tau_X}\right| \quad\Longleftrightarrow\quad \lambda_X > \lambda_M.$$

From this, it is straightforward to show that the welfare changes in (2) also continue to hold in the general case.