Trade Policy Uncertainty, Learning and Export Decision

Yuanzhe TANG*

November 13, 2021

Work in Progress

Abstract

Recent years have seen an increase in trade policy uncertainty (TPU). Conventional wisdom is that TPU deters firms' entry into export markets. However, US exports relative to GDP grew by 17 percent between 2016 and 2019 despite dramatically rising TPU. In this paper, I revisit Handley and Limão (2017) partial equilibrium model and study a firm's decision to start exporting under TPU in the presence of demand uncertainty and learning. Handley and Limão (2017) show TPU reduces entry in a sunk cost model due to Bernanke's *bad news principle* while they don't consider the difference between new and old exporters, namely age dependence. Empirical studies have suggested that export dynamics is age dependent, which is in line with the prediction of demand learning model. My goal is to verify if the result such that TPU reduces entry is robust in settings where export decision is also driven by demand learning. I first examine the effect of TPU on the timing of entry in both sunk cost learning and fixed cost learning models. More specifically, I show that a mean-preserving spread in tariff can lead to more entry if the difference between good and bad news is large enough. As in Handley and Limão (2017), bad news deters exporters' entry because it generates an option value of waiting. Moreover, in my model, good news also matters and affects early and late entry differently as early and late entrants hold different demand beliefs. The intuitive option value of waiting can be compensated by extra benefits of early learning due to early entry. I second examine the effect of variance of posterior beliefs on entry decision. The greater the variance is, the greater the benefits of learning are. As exporters become more experienced, their posterior beliefs should be closer to prior belief and they learn less from exporting. Therefore, the effect of variance of posterior beliefs sheds light on how export decision is affected by TPU as exporting age varies.

Keywords: Learning, Trade policy uncertainty

JEL Codes: D83; F12; F13

*Tang: CREST, École Polytechnique, Institut Polytechnique de Paris, 5, Av Henry Le Chatelier, 91120 Palaiseau, France (email:yuanzhe.tang@ensae.fr). I thank Geoffrey Barrows, Gregory Corcos, Hector Calvo-Pardo, Alejandro Graziano, Francis Kramarz, Olivier Loisel, Isabelle Méjean, Jean-Baptiste Michau, Julien Prat for useful comments. I am also thankful to participants at CREST seminar and 19th Annual GEP/CEPR Postgraduate Conference. All errors are mine.

1 Introduction

A growing body of empirical evidence has suggested that when making export decision, potential exporters take not only applied tariffs but also trade policy uncertainty (TPU) into account.[1] That's why during the Uruguay round, one major goal of WTO was to increase the amount of trade under binding commitments.[2] For a long time, much effort has been made to strengthen trade relationship between countries and TPU has decreased a lot especially during 90s.[3] However, in recent years, there has been a sharp increase in TPU.[4] For G20 emerging economies, the percentage of import product lines under any imposed temporary trade barriers in effect increased gradually from 0.56 percent in 1995 to 2.78 percent in 2013. In Europe, after 4 years negotiation, the UK and EU finally reached a Brexit deal at the end of 2020. Since Dec 2019, the Appellate Body has ceased functioning as the US blocked judge reappointment in order to reform WTO. In fact, even a trade war is not far away from today's international market. The intense trade dispute between the US and China has been lasting more than two years since 2018. In addition, it has been more common that trade policy is used to deal with non trade issues in recent years.[5] Even though the sharp increase in TPU over the last years is mainly driven by the threat of higher tariffs, it doesn't mean that a lower tariff world is impossible. TPP was signed during the presidency of Obama and could have been successful if Trump administration hadn't opposed the deal.[6] For the current US and China dispute, the 'Phase One' Deal was signed on January 15, 2020 and there is a possibility that a preferential trade agreement can be established after Trump's presidency.

In addition to TPU, potential exporters' entry decision can be affected by other factors. Recent literature finds that export dynamics is age dependent. Eaton, Eslava, Kugler, et al. (2007) firstly document that new exporters usually start small and exit export market quickly. Conditional on surviving, they grow fast at the beginning and become less likely to exit as export age increases. Demand learning is one of the mechanisms that can well explain findings on new exporters dynamics.

In this paper, I study a potential exporter's decision to start exporting under TPU in the presence of demand learning. The goal is to verify if the result such that TPU reduces entry is robust in models that deviate from classical sunk cost model. Common wisdom is that TPU reduces exporters' entry. In canonical models with heterogeneous firms like Melitz (2003) and Das et al.(2007), sunk entry cost matters.[7] Introducing TPU in a sunk cost model creates an option value of waiting which deters exporters' entry as pointed out by Handley and Limão (2017). However, Baley et al. (2020) argue that despite dramatically increasing policy uncertainty, US exports relative to GDP grew by 17 percent between 2016 and 2019, which raises question whether increasing TPU is an obstacle to entrants.[8] If learning drives new exporter dynamics, it might be able to encourage firms' early entry

[1] See Crowley et al. (2018), Handley (2014), Handley and Limão (2015) and Handley and Limão (2017).

[2] See a brief discussion on WTO binding commitments in Handley (2014). Using applied and bound tariffs data from 1996 to 2009, Bacchetta and Piermartini (2011) find that for bound tariff lines, the probability of an increase in applied tariffs is lower and the probability of a decrease in applied tariffs is higher.

[3] There was a massive reduction of TPU as many countries were granted access to WTO.

[4] Caldara et al. (2020) build a monthly TPU index based on the frequency of TPU terms mentioned across major newspapers and find that the average level of TPU index becomes unprecedentedly higher since 2017. Baker et al. (2016) build a US monthly TPU index based on the frequency of articles that discuss TPU related topics in over 2,000 US newspapers and the index has risen sharply in recent years. Using the same method, they find that after 2016, Brexit/EU economic uncertainty accounted for a large proportion of UK's total economic policy uncertainty.

[5] Since 2018, the Japan-South Korea trade dispute has been initiated as a response to a historical dispute over comfort women and forced labor during World War II. In 2020, president Macron stopped the negotiation on the EU-Mercosur trade agreement in order to force Brazil to deal with the severe deforestation problem in Amazon rainforest.

[6] It is still unclear whether Biden administration will reenter the agreement.

[7] See also a discussion on sunk cost, fixed cost and uncertainty in Alessandria, Arkolakis, et al. (2020).

[8] They propose a two-country general equilibrium model to explain trade promotion effect of TPU. There exists cross-border information friction which generates uncertainty about the other country's endowment. Therefore, term of trade is uncertain, which affects domestic exporting decision. Higher uncertainty can promote trade in some cases since trade can be seen as a way of risk sharing. Nevertheless, they don't consider firm level entry decision.

under TPU. From TPU perspective, by adding demand learning, I am able to explore the effect of TPU on entry with and without sunk cost. From demand learning perspective, by adding TPU, my model also sheds light on the effect of exporting age on firms' export decision under TPU.

I study both sunk cost learning and fixed cost learning models.[9] Leaning only lasts one period. Potential exporters hold prior belief. If they choose to enter, they will receive a signal at the end of their entry period and no more signals will be received after. Therefore, their posterior belief is constant after the entry period and can be either good or bad with equal probability, which depends on the signal they have received. Like Handley and Limão (2017), I assume that TPU follows a three-state Markov chain - good state, intermediate state and bad state.[10] I focus on entry cutoff in intermediate state. Entry cutoff firms are those who are indifferent between entering and waiting in the current state. In my model, there are 2 effects of TPU. One is the same as that of H&L and follows a *bad news principle* which was proposed in Bernanke (1983). Bad news discourages exporters' entry since potential entrants can prevent loss in future possible bad state by waiting. More precisely, intermediate-state entry cutoff firms prefer to wait in bad state.[11] A higher tariff in bad state only reduces the profits of early entry and makes early entry less appealing. The other effect comes from good news. If entry cutoff firms choose to enter(wait) in the current intermediate-state period, they will hold posterior(prior) belief in the next period. Therefore, future possible good news affects the value of entry and waiting differently. Combining the 2 effects above, the net effect of TPU is ambiguous.

Unlike H&L model, good news matters in my model as new exporters are different from old ones. More specifically, I show that, in sunk cost learning model, the effect of good news on intermediate-state entry cutoff is always positive. Cutoff firms are willing to enter in good state. Decreasing tariff in good state increases the profits of both early entry and waiting. Without per period fixed cost, per period export profit is always positive using CES demand and exporters keep exporting whatever the posterior belief they hold. In Lemma 2, I show that tariff reduction in good state benefits early entry relatively more and encourages firms' early entry. It is because, in per period profit function, the expectation of posterior belief related term is greater than prior belief related term. The lower the tariff in good state is, the stronger firms' willingness to start exporting and have more knowledge about their demand is. In other words, potential entrants prefer to export in future possible good state with more knowledge.

However, I show that in fixed cost learning model, the effect of good news on entry cutoff can be negative. Using fixed cost, there is endogenous exit given that export profit can be negative. In both TPU and no TPU cases, entry cutoff firms make negative profit in their entry period in order to benefit from extra profits in future better scenarios. As good state tariff is not low enough, in good state, intermediate-state entry cutoff firms make positive profit only conditional on holding good posterior belief. In this case, if good belief is not sufficiently greater than prior belief, reducing good state tariff will increase the profits of waiting relatively more and deter entry. Namely, while good posterior belief is greater than prior belief, the probability of obtaining good posterior belief is only one half, which is risky. Considering the risk of reaching the best case scenario - low tariff state and good posterior belief, and negative profit in the entry period, a late entry in future good state can be preferred if good posterior belief is not favorable enough. If good state tariff is sufficiently low, intermediate-state cutoff exporters will be able to make positive profit conditional on bad belief. In this case, reducing good state tariff increases profits of early entry more, which is similar to sunk cost learning model.

[9] Handley and Limão (H&L) use sunk cost to generate state dependence. Since demand learning also generates state dependence, I can study the effect of TPU using fixed cost only.

[10] Good state means low tariff state; intermediate state means medium tariff state and bad state means high tariff state.

[11] In fixed cost learning model, the condition such that intermediate-state cutoff exporters wait in bad state and enter in good state is not always satisfied. Some extra restrictions need to be imposed.

I explore the effect of pure uncertainty by modelling TPU as a mean-preserving spread of tariff. Enlarging the difference between bad and good state tariffs increases TPU.[12] Since per period profit is a convex function of tariff, marginal negative effect of bad news on early entry becomes smaller as bad state tariff increases. As good state tariff decreases, marginal effect of good news on early entry is positive and increasing if learning is profitable enough. As bad and good states are close to intermediate state, marginal negative effect of bad news dominates and TPU deters exporters' entry. However, as bad and good states become very different, marginal effect of good news can be positive and dominate marginal negative effect of bad news. Therefore, it's possible that a high TPU induces more entry.

I also explore the effect of the variance of posterior beliefs on entry threshold under TPU relative to no TPU case. In both TPU and no TPU cases, entry threshold is a function of posterior-beliefs variance and the relative change of these two thresholds depends on TPU process. As the variance of posterior beliefs increases and learning becomes more profitable, the effect of the variance of posterior beliefs on relative entry threshold is monotonic in sunk cost learning model but not in fixed cost learning model. Since the variance of posterior beliefs correlates negatively with exporters' age, my model is able to shed light on the impact of exporters' age on relative entry threshold under TPU.

The current paper builds on two independent literature - literature of export under TPU and literature of firm dynamics. Using a sunk cost model, Handley (2014) and Handley and Limão (2015) study the effect of TPU on entry cutoff. In a partial equilibrium, they predict that increasing TPU deters firms' early entry more.[13] Handley and Limão (2017) extend their partial equilibrium model to a general equilibrium model where aggregate price is an endogenous variable. The entry deterrence effect is not robust in some extreme cases.[14] Based on canonical Melitz model, Feng et al. (2017) assume that per-period fixed export cost is increasing in total mass of exporting firms and uncertainty reduction is equivalent to lower expectation of tariff payments. Therefore, a decrease in TPU selects more(less) productive firms into(out of) the exporting market. However, in models above, new entrants are similar to incumbents and there is no age dependence.

Canonical sunk cost model is not the only way to study the effect of TPU. Alessandria, Khan, et al. (2019) introduce TPU into a $(\underline{s}, \bar{s})$ inventory model and study the effect of yearly TPU shock on monthly import flow. Unlike H&L, they focus on the fluctuation of high frequency trade flow within each year. Using a DSGE model with endogenous customer accumulation, Steinberg (2019) disentangles numerically the cost of Brexit TPU from other macroeconomic factors. The entry threshold is analytically intractable. Conversely, my paper tries to analyse entry threshold analytically and I need to build a model as simple as possible.

Empirically, Handley (2014) studies the effect of tariff binding commitments during Uruguay Round using Australian import data and predicts that the growth of product varieties would have been 7% lower if Uruguay Round had not been implemented. In a counterfactual exercise, Handley and Limão (2015) find that if Portugal's accession to European Community had only reduced applied tariffs but not TPU, it would have achieved only 20 percent of the total predicted growth for entry and less than 30 percent for total exports. In a general equilibrium framework, Handley and Limão (2017) study the increase of China's export to US during the period of China's WTO accession and estimate an effect of reducing TPU being equivalent to a decrease in permanent tariff on Chinese

[12] Unlike Handley and Limão (2017), I use a different measure of TPU. In this paper, I also explore the effect of their TPU measure.

[13] Their model is able to give a closed form solution of entry threshold.

[14] In a general equilibrium model, if current realization of intermediate state tariff τ_1 is close to bad state tariff τ_2, TPU can encourage firms' entry. As τ_1 is close to τ_2, future tariff cannot be much higher than the current intermediate tariff and TPU process will be unfavorable to foreign domestic firms, which leads to less entry of foreign domestic firms and pushes up foreign price.

goods by 13 percentage points.[15] Using H&L framework, Crowley et al. (2018) study the indirect effect of anti-dumping duties using Chinese export data between 2000 and 2009. They assume that, within a firm, imposing anti-dumping duty on a product-market pair will generate TPU on this product and closely-related products in other markets. They also assume that policy information can be transferred across neighboring exporters. The empirical findings support their assumptions. Alessandria, Khan, et al. (2019) find that each year before the annual revision of China's MFN status, imports from China rose. However, this temporary trade increase cannot compensate for the overall trade dampening effect in the long term.

In the firm dynamics literature, demand learning is one of the mechanisms that are used to model new entrants' behavior.[16] Jovanovic (1982) studies firm dynamics by assuming that firms learn gradually and imperfectly about their unobserved type. Arkolakis et al. (2018) introduce Jovanovic's Bayesian learning in a standard monopolistically competitive environment with firm productivity heterogeneity. They find that the model predictions are consistent with empirical findings using Colombian manufacturing plant-level data. Using a fixed-effect strategy, Berman et al. (2019) show that a few empirical findings such as firm-market-specific prices and the variance of growth rates being negatively correlated with age can be explained by Bayesian demand learning while alternative demand side mechanisms fail to explain these findings. Chen et al. (2019) find direct evidence of export learning by exploiting the data on sales forecast of Japanese firms. Some other mechanisms have also been proposed to model new exporter dynamics.[17]

The literature on new exporter dynamics above hasn't formally studied entry decision under TPU while it mainly focuses on the problem of post-entry dynamics. Moreover, the literature on exporters' entry under TPU hasn't taken post-entry speciality of new exporters into consideration. This paper tries to build a bridge between the two literature where I employ a simplified demand learning mechanism to study exporter's entry decision under TPU using partial equilibrium framework.

The remainder of the paper is organized as follows. In section 2, a general setting is given. Section 3 will give a benchmark model where I study TPU using sunk cost learning model. Fixed cost learning model is presented in section 4. Section 5 gives a very preliminary discussion on my empirical application and Section 6 concludes.[18]

2 General assumptions

In this section, I will give a general setting which is applicable to both sunk cost learning and fixed cost learning models. In sunk cost learning model, per period fixed cost f is assumed to be 0 while in fixed cost learning model, sunk entry cost S is assumed to be 0. I consider a small open economy where monopolistically competitive exporters produce differentiated goods and foreign aggregate variables are taken as constant. There are infinite periods.

2.1 Consumer's demand

Representative consumer spends a fixed share of income on homogeneous good and the remaining on differentiated goods. The utility function is:

$$U_t = C_t^{\mu} Y_t^{1-\mu} \tag{1}$$

[15] Feng et al. (2017) find that after China's accession to WTO, in the US market, new Chinese exporters are more productive while those who exit are less productive.

[16] See also Albornoz et al. (2012), Timoshenko (2015a) and Timoshenko (2015b).

[17] Ruhl and Willis (2017), Foster et al. (2016), Piveteau (2016) and Fitzgerald et al. (2019) propose demand accumulation model. Aeberhardt et al. (2014) and Eaton, Eslava, Krizan, et al. (2014) propose matching and learning model.

[18] My empirical application is unfinished and there is no discussion on its results in introduction and conclusion.

Where C_t is the aggregate consumption of differentiated goods and Y_t is the aggregate consumption of homogeneous good.[19] Both are tradable. The aggregate consumption of differentiated goods is

$$C_t = \left[\int_{\omega \in \Omega_t} (e^{a_t(\omega)})^{\frac{1}{\sigma}} q_t(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right]^{\frac{\sigma}{\sigma-1}} \tag{2}$$

Where $a_t(\omega)$ is the demand shock realization of variety $\omega \in \Omega_t$ in period t. Consumer maximizes C_t such that revenue constraint $\int_{\omega \in \Omega_t} p_t(\omega) q_t(\omega) d\omega = P_t C_t \leq \mu R_t$ is satisfied. P_t is aggregate price index and R_t is aggregate revenue.[20] I assume that foreign aggregate variables are constant which are not affected by the small economy. The solution of consumer's problem is

$$q_t(\omega) = e^{a_t(\omega)} \left(\frac{p_t(\omega)}{P_t} \right)^{-\sigma} C_t \tag{3}$$

2.2 Firms' problem

Potential exporters of differentiated goods know their own productivity φ. At the beginning of each period, they can observe the realization of trade policy τ in the current period. In addition, they face an idiosyncratic demand shock. Figure 1 illustrates how demand shock evolves for a particular firm. Enter means export for the first time. If a firm enters in period t, at the beginning of period t, she will draw a true underlying demand parameter θ (θ_H or θ_L) which is unobservable and $\theta_H \geq \theta_L$. The realization of demand shock a_t in period t could be either θ_H or θ_L and its probability distribution depends on θ that firm draws. a_t is observable at the end of entry period t. p is assumed to be greater than $\frac{1}{2}$, which means the probability of $a_t = \theta_H$ is higher as $\theta = \theta_H$. By observing a_t, firm is able to form a posterior expectation about her true underlying demand parameter θ. For a firm that hasn't entered the export market, she has no demand shock to observe and holds only prior expectation.

I also assume that, for a firm that enters in period t, if she exports in period $t+1$, $t+2$, and so on, the demand shock realization will be her true underlying demand parameter θ that she has drawn in the entry period t. However, she can no longer observe her demand shock realizations from period $t+1$ to period infinite.[21]

[19] Competitive homogeneous good sector pins down the wage at unity.

[20] $P_t = \left[\int_{\omega \in \Omega_t} e^{a_t(\omega)} p_t(\omega)^{1-\sigma} d\omega \right]^{\frac{1}{1-\sigma}}$

[21] Assuming one period learning simplifies the dynamic problem as TPU is also taken into account. One-period learning cannot fully capture firm's learning dynamics across different periods while it can still give some insights on how demand learning and TPU jointly affect firm's entry decision. Assuming demand shock realizations from period $t+1$ to period infinite are no longer observed is equivalent to assuming that firm no longer observes her export revenue after period t.

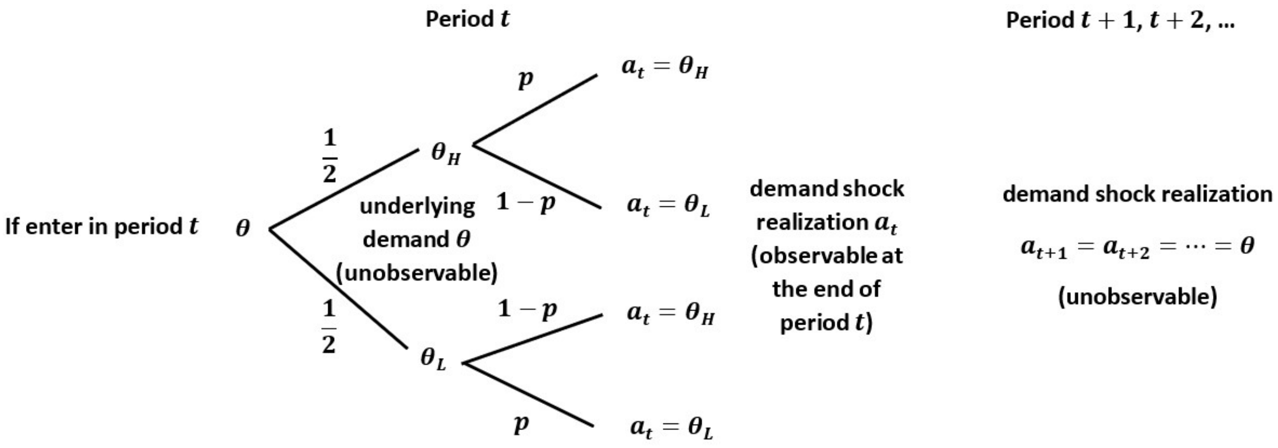

Fig. 1: Demand shock

The TPU process follows a three-state Markov chain. If the current state is high tariff state τ_2 (low tariff state τ_0), the state in the next period will also be τ_2 (τ_0). If the current state is intermediate tariff state τ_1, with probability $1-\gamma$, the state will be τ_1 in the next period and with probability $\gamma\lambda_2$ ($\gamma(1-\lambda_2)$), the state will be τ_2 (τ_0) in the next period. I assume that $\tau_2 \geq \tau_1 \geq \tau_0$. τ_2 is bad news and τ_0 is good news. I focus on firm's entry decision in intermediate state τ_1 and also study the effect of pure uncertainty by assuming that τ_2 and τ_0 is a mean-preserving spread of τ_1 and $\tau_1 = \lambda_2\tau_2 + (1-\lambda_2)\tau_0$. The mean-preserving spread can be rewritten as $\tau_2 = \delta\tau_1$ and $\tau_0 = \frac{1-\lambda_2\delta}{1-\lambda_2}\tau_1$ with $\delta \geq 1$ and $\tau_0 \geq 1$. Figure 2 illustrates the TPU process.

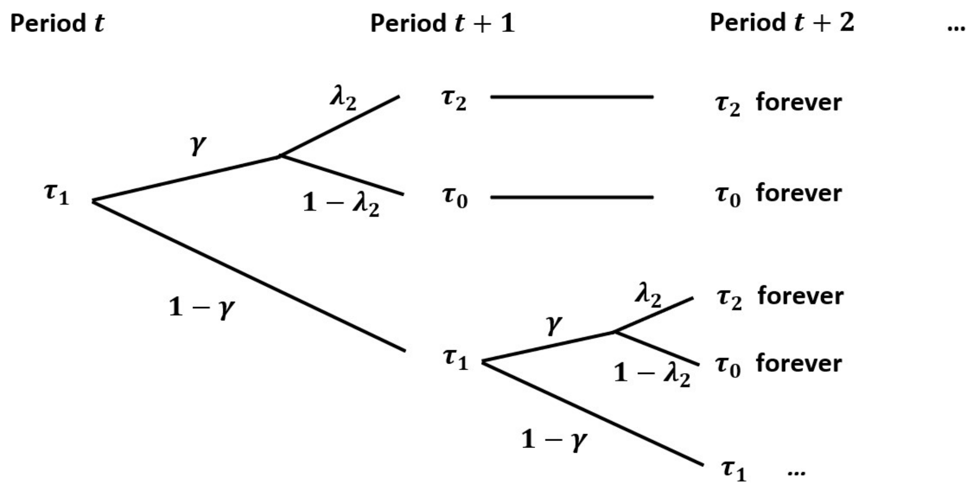

Fig. 2: TPU process

As a firm decides to export in period t, the quantity of goods being exported doesn't affect her learning about demand shock at the end of period t. Therefore, once a firm chooses to export in period t, she produces a quantity such that the expected profit of current period is maximized. For an exporter with productivity φ that exports in period t, her conditional expected per period profit in period t will be

$$E_t\pi_t(\varphi, \tau_t, \bar{a}_t) = E_t\left(\frac{p_t q_t}{\tau_t} - \frac{q_t}{\varphi} - f|\varphi, \tau_t, \bar{a}_t\right) \tag{4}$$

Where $\tau_t \geq 1$ is the realization of ad valorem tariff in period t which is observable at the beginning of period t and $\bar{a}_t$ is firm's past demand shock realization.[22] f is per period fixed cost. If period t is firm's entry period, she will make production decision without any information about her underlying demand parameter θ at the beginning of period t. If her entry period is period $t-i$ with $i \geq 1$, in period t, the firm will make production decision based on the signal $\bar{a}_t = a_{t-i}$ she has received in entry period $t-i$.[23] The exporter in period t maximizes $E_t\pi_t$ by choosing production quantity q_t conditional on her private information φ, $\bar{a}_t$ and public information τ_t given that $p_t = \left(\frac{e^{a_t}\mu R}{q_t}\right)^{\frac{1}{\sigma}} P^{\frac{\sigma-1}{\sigma}}$.[24] The quantity being chosen is

$$q_t = \left(\frac{\sigma-1}{\sigma}\right)^{\sigma}\left(\frac{b_t\varphi}{\tau_t}\right)^{\sigma}\frac{\mu R}{P^{1-\sigma}} \tag{5}$$

Where b_t is firm's belief about her demand shock realization in period t. More specifically, if $t-i$ is entry period and $\bar{a}_t = a_{t-i} = \theta_H$, firm's belief in period t will be a good posterior belief $b_t(\bar{a}_t = \theta_H) = E_t\left(e^{\frac{a_t}{\sigma}}|\bar{a}_t = \theta_H\right) = b_H$. If $\bar{a}_t = a_{t-i} = \theta_L$, firm's belief in period t will be a bad posterior belief $b_t(\bar{a}_t = \theta_L) = E_t\left(e^{\frac{a_t}{\sigma}}|\bar{a}_t = \theta_L\right) = b_L$. If t is entry period, firm's belief in period t will be a prior belief $b_t = E_t(e^{\frac{a_t}{\sigma}}) = b_M$ which is an unconditional expectation. Firms don't forget their posterior belief even if they stop exporting. As shown in Appendix A, $b_H \geq b_M \geq b_L > 0$ and $b_H + b_L = 2b_M$, which can be rewritten as $b_H = \varepsilon b_M$ and $b_L = (2-\varepsilon)b_M$ with $\varepsilon \in [1,2)$. Figure 3 illustrates the belief process where $\frac{1}{2}$ is the unconditional probability such that belief is high b_H(low b_L). Bring the quantity decision back into expected profit function and the solution of per period expected profit of exporting is

$$E_t\pi_t(\varphi,\tau_t,\bar{a}_t) = \frac{(\sigma-1)^{\sigma-1}}{\sigma^{\sigma}}b_t^{\sigma}\tau_t^{-\sigma}\varphi^{\sigma-1}\frac{\mu R}{P^{1-\sigma}} - f \tag{6}$$

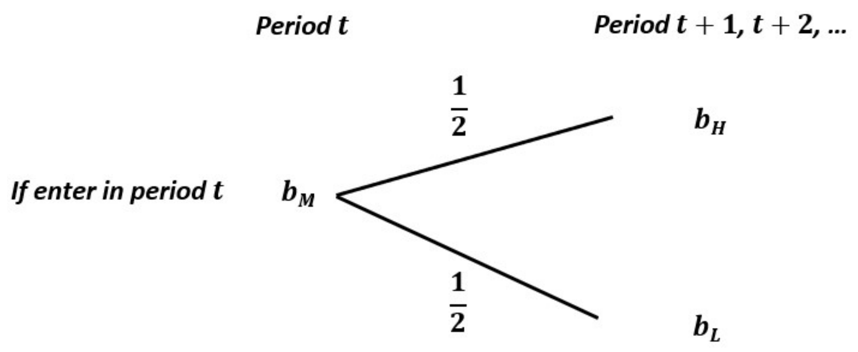

Fig. 3: Demand belief

Therefore, the solution of exporter's conditional per period expected profit depends on her productivity φ, current trade policy realization τ and her current belief b. Equation (6) can be rewritten as

$$\pi(\varphi,\tau,b) = b^{\sigma}\tau^{-\sigma}\varphi^{\sigma-1}k - f \tag{7}$$

[22] The result of iceberg cost is very similar to that of ad valorem tariff.

[23] Because I assume that learning process only lasts one period. If period t is entry period, $\bar{a}_t$ is empty. If period $t-i$ is entry period, $\bar{a}_t = a_{t-i}$.

[24] As foreign aggregate variables are assumed to be constant, $R_t = R$ and $P_t = P$. Learning lasts one period and only the entry period price can be observed at the end of entry period, which means firms can only observe their export revenue of entry period.

Where $k = \frac{(\sigma-1)^{\sigma-1}}{\sigma^\sigma}\frac{\mu R}{P^{1-\sigma}}$ is a constant. Recall that the unconditional probability of having a good belief b_H (bad belief b_L) is $\frac{1}{2}$ and $b_H + b_L = 2b_M$. There is $\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma \geq b_M^\sigma$ as $\sigma > 1$. Unconditionally, exporter prefers knowing the signal to not knowing the signal, which is equivalent to saying that exporter prefers more knowledge about her true underlying demand parameter.[25]

3 Sunk cost and learning

In this section, I introduce one-period demand learning in the partial equilibrium model of Handley and Limão (2017) where potential exporters pay sunk cost S to enter and there is no per period fixed cost ($f = 0$). Table 1 compares H&L partial equilibrium model with my sunk cost learning model. If there is no demand learning and $b_H = b_L = b_M$, sunk cost learning model will be equivalent to H&L partial equilibrium model. Therefore, H&L partial equilibrium model can be seen as a special case of my model. However, as there is no fixed cost, neither models can capture endogenous exit.[26] I will firstly give the benchmark entry threshold φ_1 in a no TPU case and then consider a case with TPU.

	H&L (2017)	sunk cost learning
sunk entry cost	$S > 0$	$S > 0$
fixed cost	$f = 0$	$f = 0$
endogenous exit	no	no
expected per period profit	$b_M^\sigma \tau^{-\sigma}\varphi^{\sigma-1}k$	$b^\sigma \tau^{-\sigma}\varphi^{\sigma-1}k$ b_M if entry period; b_H or b_L otherwise
solution of $\frac{\varphi_{1u}}{\varphi_1}$ relative entry threshold under TPU	explicit	explicit
TPU deters entry only $\frac{\varphi_{1u}}{\varphi_1} \geq 1$	yes	depends

Tab. 1: Difference between H&L (2017) and sunk cost learning model

3.1 No TPU case

Define φ_1 as the entry threshold in a case where there is no TPU and tariff is constant τ_1. Define $\Pi_e(\varphi, \tau)$ as the expected value from exporting after entry (after paying sunk cost) with entry condition being τ and φ. If there is no TPU, for potential exporters, entry condition will be the same in each period. Therefore, waiting cannot bring extra profits and entry cutoff firms are those whose expected value from exporting after entry equals to sunk entry cost S. We have

$$\Pi_e(\varphi_1, \tau_1) = \pi(\varphi_1, \tau_1, b_M) + \frac{1}{2}\frac{\beta}{1-\beta}\Big[\pi(\varphi_1, \tau_1, b_H) + \pi(\varphi_1, \tau_1, b_L)\Big] = S \tag{8}$$

$$\Leftrightarrow \varphi_1^{\sigma-1} = \frac{S\tau_1^\sigma}{\Big[b_M^\sigma + \frac{1}{2}\frac{\beta}{1-\beta}\left(b_H^\sigma + b_L^\sigma\right)\Big]k} \tag{9}$$

$$\Rightarrow \varphi_1^{*\sigma-1}\Big|_{b_H=b_L=b_M} = \frac{S\tau_1^\sigma}{\Big[b_M^\sigma + \frac{\beta}{1-\beta}b_M^\sigma\Big]k} \tag{10}$$

[25] The inequation I have in a demand learning model is similar to that in a passive customer base accumulation model.

[26] In the latter section, I will show that fixed cost complicates the model much. Therefore, I don't further study a model where both sunk cost and fixed cost are included.

Where φ_1^* is H&L entry threshold without TPU. Recall that $\frac{1}{2}\left(b_H^\sigma + b_L^\sigma\right) \geq b_M^\sigma$. Therefore, $\varphi_1 \leq \varphi_1^*$ and demand learning can bring extra entrants. Since $b_H = \varepsilon b_M$ and $b_L = (2-\varepsilon) b_M$, as the variance measure of posterior beliefs ε increases, φ_1 decreases and there will be more entrants.[27] The higher the variance measure of belief ε is, the more profitable learning is.

3.2 TPU case

In this section, I consider a case with TPU. The expected value from exporting after entry $\Pi_e(\varphi, \tau)$ can be rewritten using a recursive formula.[28]

$$\Pi_e(\varphi, \tau) = f(\varphi, \tau) + \beta E_\tau \Pi_e(\varphi, \tau') \tag{11}$$

Where τ is the current trade policy realization and τ' is the trade policy realization in the next period. Unlike H&L (2017), $f(\varphi, \tau)$ is no longer the current per period profit $\pi(\varphi, \tau, b)$. The belief in the entry period b_M is different from those in post-entry periods (b_H or b_L) and this difference will be taken into account when writing $f(\varphi, \tau)$. Therefore, in $f(\varphi, \tau)$, factor $\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma - b_M^\sigma$ appears. Define $\Pi(\varphi, \tau)$ as the expected value under trade policy τ for a firm φ.

$$\Pi(\varphi, \tau) = \max\left\{\Pi_e(\varphi, \tau) - S, \beta E_\tau \Pi(\varphi, \tau')\right\} \tag{12}$$

$\Pi(\varphi, \tau)$ is the maximum between entering in the current period $\Pi_e(\varphi, \tau) - S$ and waiting in the current period $\beta E_\tau \Pi(\varphi, \tau')$. Minus each side by $\Pi_e(\varphi, \tau) - S$ in (12) and bring (11) into (12).

$$\Pi(\varphi, \tau) - \Pi_e(\varphi, \tau) + S = \max\left\{0, \beta E_\tau\left[\Pi(\varphi, \tau') - \Pi_e(\varphi, \tau') + S\right] - f(\varphi, \tau) + (1-\beta) S\right\} \tag{13}$$

$\Pi(\varphi, \tau) - \Pi_e(\varphi, \tau) + S$ is the value net of the profits of entering in the current period. If it's positive, firm φ will choose to wait in the current period. If it's 0, firm φ will enter in the current period. Define $V(\varphi, \tau) = \Pi(\varphi, \tau) - \Pi_e(\varphi, \tau) + S$ as the value of waiting

$$V(\varphi, \tau) = \max\left\{0, \beta E_\tau V(\varphi, \tau') - f(\varphi, \tau) + (1-\beta) S\right\} \tag{14}$$

I am only interested in the entry threshold in intermediate tariff state τ_1 - φ_{1u}. φ_{1u} firms are indifferent between entering and waiting under τ_1. Therefore, the following condition should be satisfied.

$$\beta E_{\tau_1} V(\varphi_{1u}, \tau') - f(\varphi_{1u}, \tau_1) + (1-\beta) S = 0 \tag{15}$$

Equation (15) implies that $V(\varphi_{1u}, \tau_1) = 0$. $V(\varphi_{1u}, \tau_0) = 0$ and $V(\varphi_{1u}, \tau_2) > 0$ are also satisfied. φ_{1u} firm is willing to enter in low tariff state τ_0 and wait in high tariff state τ_2.[29] The explicit solution of $\frac{\varphi_{1u}}{\varphi_1}$ is[30]

$$\frac{\varphi_{1u}^{\sigma-1}}{\varphi_1^{\sigma-1}} = \frac{1 + \frac{1}{2}\beta\left[b(\varepsilon) - 2\right] + \beta\gamma\lambda_2\left(1 + \frac{1}{2}\frac{\beta}{1-\beta}b(\varepsilon)\right)}{1 + \frac{1}{2}\beta\left[b(\varepsilon) - 2\right]\left[1 - \gamma + \gamma\lambda_2\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma} + \gamma\left(1-\lambda_2\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\right] + \beta\gamma\lambda_2\left(1 + \frac{1}{2}\frac{\beta}{1-\beta}b(\varepsilon)\right)\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}} \tag{16}$$

[27] ε is not the real variance of posterior beliefs but they are positively correlated. In this paper, I just call ε the variance measure of posterior beliefs.

[28] See Appendix B for more details.

[29] See more discussion about value function $V(\varphi, \tau)$ in Appendix C

[30] $\frac{\varphi_{1u}}{\varphi_1}$ is the relative entry threshold under TPU. See Appendix D for more details.

Where $b(\varepsilon) = \varepsilon^\sigma + (2-\varepsilon)^\sigma$ and $b(\varepsilon) - 2 \geq 0$. Good belief b_H is substituted by εb_M and bad belief b_L is substituted by $(2-\varepsilon)b_M$. In the above equation, all b_M terms cancel out. As $\varepsilon = 1$, $b_H = b_L = b_M$ and we have H&L solution $\frac{{\varphi^*}_{1u}^{\sigma-1}}{{\varphi^*}_{1}^{\sigma-1}}$.

$$\left.\frac{{\varphi^*}_{1u}^{\sigma-1}}{{\varphi^*}_{1}^{\sigma-1}}\right|_{\varepsilon=1} = \frac{1+\frac{\beta}{1-\beta}\gamma\lambda_2}{1+\frac{\beta}{1-\beta}\gamma\lambda_2\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}} \geq 1 \tag{17}$$

In H&L, $\varphi^*_{1u} \geq \varphi^*_1$. Firm's entry threshold is greater under TPU and TPU only deters firm's entry. Moreover, the negative effect of TPU only comes from high tariff τ_2, which is called *bad news principle*. However, in equation (16), low tariff state τ_0 also affects the solution of $\frac{\varphi_{1u}}{\varphi_1}$. Good news τ_0 matters as learning is introduced since incumbents and potential exporters hold different beliefs. If there is no TPU and $\tau_2 = \tau_0 = \tau_1$, the solution of $\frac{\varphi_{1u}}{\varphi_1}$ will be equal to 1.

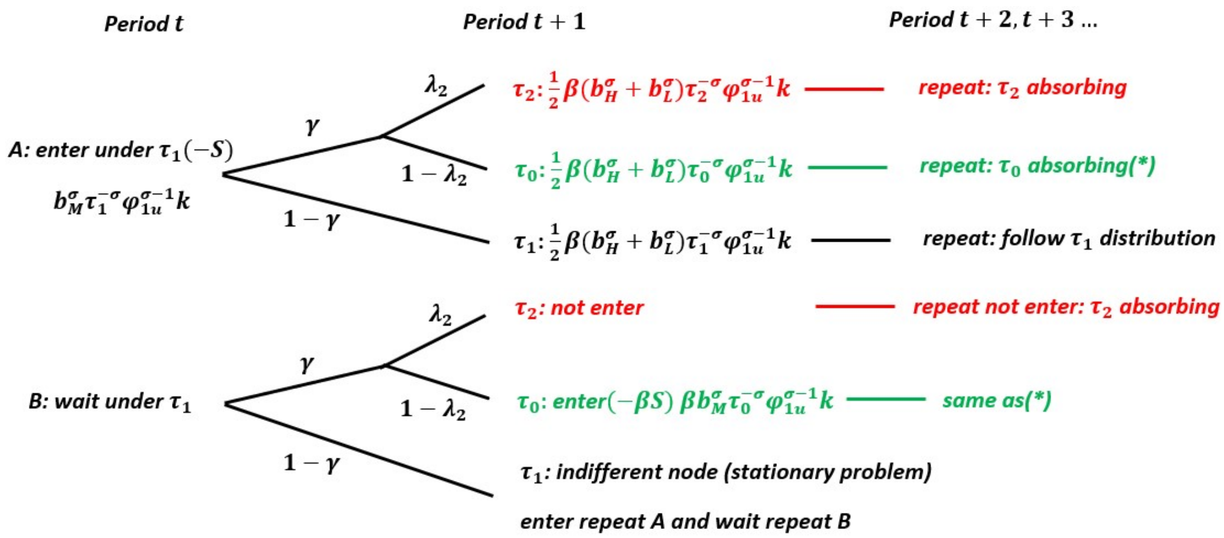

Fig. 4: Sunk cost decision tree

Figure 4 shows the decision tree of entry cutoff firms φ_{1u}. φ_{1u} firms are indifferent between entering and waiting in intermediate tariff state τ_1. The red terms are those related to bad news τ_2 and the green terms are those related to good news τ_0.[31] There are two following lemmas.

Lemma 1

Bad news principal: increasing bad state tariff τ_2 increases $\frac{\varphi_{1u}}{\varphi_1}$ and deters firms' early entry under TPU.

Proof: from equation (16), it's easy to verify that $\frac{\varphi_{1u}}{\varphi_1}$ increases in τ_2.

The intuition of bad news principal is simple. Let us focus on the red terms in Figure 4. Since entry cutoff firms φ_{1u} are willing to wait under τ_2, increasing bad state tariff τ_2 has no effect on the

[31] In fact, there are 2 black terms that are also related to τ_2 and τ_0. One is *repeat: follow τ_1 distribution* and the other is τ_1*: indifferent node*. However, if we keep drawing these 2 terms, we can find that they just repeat the red and green terms. If, in indifferent node, option A is chosen, the two black terms will cancel out from period $t+2$. If, in indifferent node, option B is chosen, then from period $t+2$, the two black terms will just repeat the same thing as that in period $t+1$. Therefore, the red and green terms are able to capture the sign of marginal effect of τ_2 and τ_0 on entry threshold φ_{1u}.

value of red term in option B - waiting under τ_1. However, a higher τ_2 will reduce the value of red term in option A - entering under τ_1 directly. Therefore, future possible bad state creates an option value of waiting and a higher bad state tariff only deters firms' early entry more. It is also easy to verify that as the probability of reaching bad tariff state λ_2 increases, $\frac{\varphi_{1u}}{\varphi_1}$ increases and there are less entrants under TPU.[32]

Lemma 2

Good news principal: decreasing good state tariff τ_0 decreases $\frac{\varphi_{1u}}{\varphi_1}$ and encourages firms' early entry under TPU.

Proof: from equation (16), it's easy to verify that $\frac{\varphi_{1u}}{\varphi_1}$ increases in τ_0.

Now let us focus on the green terms in Figure 4. Since entry cutoff firms φ_{1u} are willing to enter in low tariff state τ_0, decreasing good state tariff τ_0 increases value of green terms in both option A and option B - entering and waiting under τ_1 simultaneously. Recall that $b_H^\sigma + b_L^\sigma \geq b_M^\sigma$. Learning can bring extra profits and entering early means learning early. Therefore, the effect of decreasing τ_0 is magnified by early learning and decreasing τ_0 encourages firms' early entry more. In other words, firms prefer to have more knowledge about their underlying demand parameter as good news comes.[33] In order to capture the pure effect of uncertainty, I consider a special case where TPU is a mean-preserving spread (MPS). Substitute τ_2 by $\delta\tau_1$ and τ_0 by $\frac{1-\lambda_2\delta}{1-\lambda_2}\tau_1$. The solution of $\frac{\varphi_{1u}}{\varphi_1}$ is

$$\left.\frac{\varphi_{1u}^{\sigma-1}}{\varphi_1^{\sigma-1}}\right|_{MPS} = \frac{1+\frac{1}{2}\beta\left[b(\varepsilon)-2\right]+\beta\gamma\lambda_2\left(1+\frac{1}{2}\frac{\beta}{1-\beta}b(\varepsilon)\right)}{1+\frac{1}{2}\beta\left[b(\varepsilon)-2\right]\left[1-\gamma+\gamma\lambda_2\delta^{-\sigma}+\gamma\left(1-\lambda_2\right)\left(\frac{1-\lambda_2\delta}{1-\lambda_2}\right)^{-\sigma}\right]+\beta\gamma\lambda_2\left(1+\frac{1}{2}\frac{\beta}{1-\beta}b(\varepsilon)\right)\delta^{-\sigma}} \tag{18}$$

δ measures the distance between high tariff τ_2 and intermediate tariff τ_1. The larger δ is, the higher the variance of tariff shock is. Therefore, δ is one of the TPU measures. There is a following proposition.

Proposition 1

As the level of TPU increases from zero TPU - $\delta = 1$, $\frac{\varphi_{1u}}{\varphi_1}$ increases firstly and there are less entrants. However, $\frac{\varphi_{1u}}{\varphi_1}$ may start to decrease if δ is sufficiently large and there can be more entrants.

Proof: instead of taking the first and second derivative of $\frac{\varphi_{1u}^{\sigma-1}}{\varphi_1^{\sigma-1}}$ regarding to δ, take the first and second derivative of $\frac{\varphi_1^{\sigma-1}}{\varphi_{1u}^{\sigma-1}}$. It is easy to verify that $\frac{\partial^2 \frac{\varphi_1^{\sigma-1}}{\varphi_{1u}^{\sigma-1}}}{\partial\delta^2} > 0$ (convexity) and $\left.\frac{\partial \frac{\varphi_1^{\sigma-1}}{\varphi_{1u}^{\sigma-1}}}{\partial\delta}\right|_{\delta=1} < 0$.

Recall that, bad news τ_2 only affects red terms in option A - enter under τ_1 while good news τ_0 affects green terms in both option A - enter under τ_1 and option B - wait under τ_1. Therefore, the net negative effect of bad news on early entry is for multiple periods in bad state while the net

[32] See Appendix E for more details.

[33] It will be easier to understand good news principal if we consider demand learning as another form of passive customer base accumulation. When good news comes, firms always prefer to have a larger customer base at the same time.

positive effect of good news on early entry is just for one period in good state.[34] Also recall that per period profit $\pi(\varphi,\tau,b)=b^{\sigma}\tau^{-\sigma}\varphi^{\sigma-1}k$ is a convex function of tariff τ.[35] As τ increases, the marginal negative effect of τ on $\pi(\varphi,\tau,b)$ decreases. The overall marginal effect of TPU measure δ on early entry through bad and good news channels jointly depends on the number of periods being influenced and the magnitude of marginal effect of τ on per period profit. As $\delta=1$, $\tau_2=\tau_0=\tau_1$ and there is no TPU. In this case, it is obvious that the entry threshold under TPU - φ_{1u} is equal to the entry threshold under trade policy certainty - φ_1. As a small TPU δ is imposed, bad state tariff $\tau_2=\delta\tau_1$ increases slightly above τ_1 and good state tariff $\tau_0=\frac{1-\lambda_2\delta}{1-\lambda_2}\tau_1$ decreases slightly below τ_1. In this case, the magnitude of marginal effect through bad news and good news on per period profit is similar as the probability of reaching both states is considered. However, since more periods in bad state are influenced by bad news, bad news effect dominates good news effect and there are less entrants under TPU. As δ keeps increasing and the gap between bad news τ_2 and good news τ_0 becomes larger, the difference between marginal negative effect through bad news and marginal positive effect through good news on per period profit increases. For a sufficiently large δ, it is possible that the one-period positive effect through good news dominates multiple-periods negative effect through bad news. In this case, $\frac{\varphi_{1u}}{\varphi_1}$ may decreases below one and TPU can encourage firms' early entry.

H&L (2017) use the probability of a tariff change γ as a measure of TPU. In their partial equilibrium model, as γ increases, there are less entrants under TPU because of *bad news principal*. However, in my model, γ cannot determine if the effect of TPU on entry cutoff is negative or not. There is a following proposition.

Proposition 2

The probability of a tariff change γ only affects the magnitude of TPU effect but not the sign (positive or negative) of TPU effect.

Proof: by taking the first derivative of $\frac{\varphi_{1u}}{\varphi_1}$ regarding to γ, it is easy to verify that $\frac{\partial\frac{\varphi_{1u}}{\varphi_1}}{\partial\gamma}\geq 0$ if and only if $\frac{\varphi_{1u}}{\varphi_1}\geq 1$ and $\frac{\partial\frac{\varphi_{1u}}{\varphi_1}}{\partial\gamma}\leq 0$ if and only if $\frac{\varphi_{1u}}{\varphi_1}\leq 1$. See Appendix F for more details.

The intuition is straightforward. γ is the probability of a tariff change and it doesn't alter the shape of tariff shock. Only λ_2, τ_2 and τ_0 can alter the shape of tariff shock.[36] Therefore, γ cannot determine if the net effect of TPU is negative ($\frac{\varphi_{1u}}{\varphi_1}\geq 1$) or not ($\frac{\varphi_{1u}}{\varphi_1}\leq 1$). And a higher γ only magnifies the net effect of TPU since there is a higher probability of tariff shock hitting.

I also consider effect of the variance measure of posterior beliefs ε. The expectation of posterior belief related term $\frac{1}{2}b_H^{\sigma}+\frac{1}{2}b_L^{\sigma}=\frac{1}{2}\left(\varepsilon b_M\right)^{\sigma}+\frac{1}{2}\left(\left(2-\varepsilon\right)b_M\right)^{\sigma}$ is an increasing function of ε and higher ε means higher profits of learning. A higher ε makes option A - early entry more appealing because of early learning and the entry cutoff under TPU φ_{1u} will decrease. However, as ε increases, from equation (9), the entry threshold in a no TPU case φ_1 will also decrease.[37] Therefore, $\frac{\varphi_{1u}}{\varphi_1}$ can capture the relative change of the 2 entry thresholds and there is a following proposition.

Proposition 3

The relative entry threshold $\frac{\varphi_{1u}}{\varphi_1}$ is increasing in ε if and only if $\lambda_2\tau_2^{-\sigma}+(1-\lambda_2)\tau_0^{-\sigma}<\tau_1^{-\sigma}$ and

[34] It is because I assume that learning only lasts one period.

[35] The convexity still holds using linear demand. It also holds using iceberg cost as the elasticity of substitution σ is greater than 1. Part of the convexity is due to the assumption of ad valorem tariff since for given price p and quantity q, firm's revenue $\frac{pq}{\tau}$ is a convex function of τ.

[36] γ can be seen as a common factor of tariff shock. As γ increases, the probability of reaching high tariff state τ_2 and low tariff state τ_0 increases proportionally.

[37] Recall that TPU related parameters - τ_2, τ_0, λ_2, γ and δ only affect φ_{1u} but not φ_1.

decreasing in ε otherwise. A TPU process such that $\lambda_2\tau_2^{-\sigma} + (1-\lambda_2)\tau_0^{-\sigma} \le \tau_1^{-\sigma}$ can never bring relatively more entrants under TPU.

Proof: take the first derivative of $\frac{\varphi_{1u}}{\varphi_1}$ regarding to ε and its monotonicity depends on the relation between $\lambda_2\tau_2^{-\sigma} + (1-\lambda_2)\tau_0^{-\sigma}$ and $\tau_1^{-\sigma}$. As $\varepsilon = 1$, $\frac{\varphi_{1u}}{\varphi_1}$ is H&L solution which is greater than 1. As $\lambda_2\tau_2^{-\sigma} + (1-\lambda_2)\tau_0^{-\sigma} \le \tau_1^{-\sigma}$, $\frac{\varphi_{1u}}{\varphi_1}$ is a non-decreasing function of ε and φ_{1u} is always greater than φ_1. See Appendix G for more details.

Recall that per period profit is $\pi(\varphi,\tau,b) = b^\sigma\tau^{-\sigma}\varphi^{\sigma-1}k$. Therefore, $\lambda_2\tau_2^{-\sigma} + (1-\lambda_2)\tau_0^{-\sigma}$ can be seen as an inverted weighted average of tariff shock. $\lambda_2\tau_2^{-\sigma} + (1-\lambda_2)\tau_0^{-\sigma} < \tau_1^{-\sigma}$ means that the imposed tariff shock is less favorable than the intermediate tariff τ_1. In this case, even though a higher ε decreases both φ_{1u} and φ_1, φ_1 decreases more since the profits of learning are proportional to tariff. In order to have relatively more entrants under TPU, there should be a TPU process favorable enough such that $\lambda_2\tau_2^{-\sigma} + (1-\lambda_2)\tau_0^{-\sigma} > \tau_1^{-\sigma}$.[38] Similarly, as we increase intermediate tariff τ_1, both φ_{1u} and φ_1 increase. Moreover, it is easy to verify that $\frac{\varphi_{1u}}{\varphi_1}$ is decreasing in τ_1 from equation (16). Compared to no TPU case, a higher intermediate tariff τ_1 deters firms' early entry relatively less under TPU. As τ_1 increases, both τ_2 and τ_0 become relatively lower. Therefore, bad news effect is relatively weaker and good news effect is relatively stronger. The effect of TPU becomes relatively more positive, which makes the cost of increasing τ_1 relatively smaller under TPU.

In this section, I introduce one-period learning in H&L partial equilibrium model using sunk cost. Like H&L, I obtain a closed form solution and H&L model is nested in my new model. However, using only sunk cost, there is no endogenous exit. In next section, I will use per period fixed cost instead of sunk cost and also focus on the entry decision under TPU. Unfortunately, using fixed cost, the model becomes less tractable.

4 Fixed cost and learning

Here, I study a model with fixed cost and learning but without sunk cost. Using fixed cost, the per period expected profit of exporting $\pi(\varphi,\tau,b)$ can be negative. Facing a negative expected profit, exporter may choose to exit the export market endogenously. Table 2 compares fixed cost learning model with sunk cost learning model and H&L model. One major difference here is that there is no easy solution of relative entry threshold $\frac{\varphi_{1u}}{\varphi_1}$. In addition, the marginal effect of good news τ_0 can be negative.

	H&L (2017)	sunk cost learning	fixed cost learning
sunk entry cost	$S > 0$	$S > 0$	$S = 0$
fixed cost	$f = 0$	$f = 0$	$f > 0$
endogenous exit	no	no	yes
expected per period profit	$b_M^\sigma\tau^{-\sigma}\varphi^{\sigma-1}k$	$b^\sigma\tau^{-\sigma}\varphi^{\sigma-1}k$ b_M if entry period; b_H or b_L otherwise	$b^\sigma\tau^{-\sigma}\varphi^{\sigma-1}k - f$ b_M if entry period; b_H or b_L otherwise
solution of $\frac{\varphi_{1u}}{\varphi_1}$ relative entry threshold under TPU	explicit	explicit	implicit
TPU deters entry only $\frac{\varphi_{1u}}{\varphi_1} \ge 1$	yes	depends	depends

Tab. 2: Difference between 3 models

[38] Using a mean-preserving spread, $\lambda_2\tau_2^{-\sigma} + (1-\lambda_2)\tau_0^{-\sigma} > \tau_1^{-\sigma}$ is satisfied.

4.1 No TPU case

Recall that φ_1 is the entry threshold when there is no TPU and tariff is constant τ_1. $\Pi_e(\varphi,\tau)$ is the expected value of exporting with entry condition being τ and φ. As before, since there is no TPU, entry condition is the same across each period and waiting cannot bring extra profits. Therefore, entry cutoff firms are those whose expected value of exporting equals to 0. We have

$$\Pi_e(\varphi_1,\tau_1)=\pi(\varphi_1,\tau_1,b_M)+\frac{1}{2}\frac{\beta}{1-\beta}\Bigg(\max\{\pi(\varphi_1,\tau_1,b_H),0\}+\max\{\pi(\varphi_1,\tau_1,b_L)\}\Bigg)=0 \tag{19}$$

$$\Leftrightarrow \pi(\varphi_1,\tau_1,b_M)+\frac{1}{2}\frac{\beta}{1-\beta}\pi(\varphi_1,\tau_1,b_H)=0 \tag{20}$$

$$\Leftrightarrow \varphi_1^{\sigma-1}=\frac{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)f}{\left(b_M^\sigma+\frac{1}{2}\frac{\beta}{1-\beta}b_H^\sigma\right)k}\tau_1^\sigma \tag{21}$$

$$\Rightarrow \varphi_1^{*\sigma-1}|_{b_H=b_M=b_L}=\frac{f}{b_M^\sigma k}\tau_1^\sigma \tag{22}$$

The left hand side of equation (19) gives the expression of expected value of exporting. Since $\pi(\varphi,\tau,b_L)\le\pi(\varphi,\tau,b_M)\le\pi(\varphi,\tau,b_H)$, equation (19) implies that $\pi(\varphi_1,\tau_1,b_L)\le\pi(\varphi_1,\tau_1,b_M)\le 0\le \pi(\varphi_1,\tau_1,b_H)$ should be satisfied. Hence, equation (19) can be rewritten as equation (20). In order to benefit from future positive expected profits conditional on receiving a good signal and having good belief b_H, entry cutoff firms φ_1 are willing to suffer from a negative expected profit in the entry period. φ_1^* is the entry cutoff without demand learning which is also the zero per-period profit cutoff conditional on prior belief b_M.[39] $\varphi_1<\varphi_1^*$ and learning encourages firms' early entry. Recall that $b_H=\varepsilon b_M$. Like sunk cost learning model, as the variance measure of belief ε increases, φ_1 decreases and there will be more entrants.

In sunk cost learning model, without per period fixed cost, per period profit is always positive whatever the scenario is and entry firms take both good and bad beliefs into account. However, using fixed cost, exporters won't export with a negative per period profit if there is no more learning. In other words, for post-learning periods, entry firms take only the scenarios with positive profits into consideration. Because of free exit, loosely speaking, entry firms are willing to take more risks. In the current case, φ_1 firms only care about how good $\pi(\varphi_1,\tau_1,b_H)$ is while they don't care about how bad $\pi(\varphi_1,\tau_1,b_L)$ is. This difference will also affect the property of good and bad news under TPU.

4.2 TPU case

Recall that, in sunk cost learning model, intermediate-state entry cutoff firms φ_{1u} are willing to enter in low tariff state τ_0 but wait in high tariff state τ_2. However, using fixed cost, this condition may not hold and some extra assumption needs to be imposed. In the rest of the section, I assume that φ_{1u} firms prefer to enter under τ_0 and wait under τ_2, which seems to be a reasonable assumption. Following similar steps as those in sunk cost learning model, the implicit solution of $\frac{\varphi_{1u}}{\varphi_1}$ is[40]

$$\frac{\varphi_{1u}^{\sigma-1}}{\varphi_1^{\sigma-1}}=\frac{numerator}{denominator} \tag{23}$$

[39] $\pi(\varphi_1^*,\tau_1,b_M)=0$ and φ_1^* is the zero per-period profit cutoff.

[40] See Appendix H for more details. In Appendix H, I also show that $\pi(\varphi_{1u},\tau_1,b_M)<0$. φ_{1u} firms make negative profit in the entry period in order to benefit from future possible higher profits because of demand learning.

$$
\begin{aligned}
numerator = &\Bigg(1-(1-\gamma)\,\beta\left(-\frac{1}{2}\mathbb{1}_{\varphi_{1u},\tau_1,b_H}+1\right)-\gamma\lambda_2\beta\left(-\frac{1}{2}\mathbb{1}_{\varphi_{1u},\tau_2,b_H}+1\right)\\
&-\gamma(1-\lambda_2)\,\beta\left(\frac{1}{2}-\frac{1}{2}\mathbb{1}_{\varphi_{1u},\tau_0,b_L}\right)+\frac{\beta\gamma\lambda_2}{1-\beta}-\frac{\beta\gamma\lambda_2}{1-\beta}\beta\left(-\frac{1}{2}\mathbb{1}_{\varphi_{1u},\tau_2,b_H}+1\right)\Bigg)\frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^{\sigma}}{1+\frac{1}{2}\frac{\beta}{1-\beta}}
\end{aligned} \tag{24}
$$

$$
\begin{aligned}
denominator = &1+(1-\gamma)\,\beta\left(\frac{1}{2}\mathbb{1}_{\varphi_{1u},\tau_1,b_H}\varepsilon^{\sigma}-1\right)+\gamma\lambda_2\beta\left(\frac{1}{2}\mathbb{1}_{\varphi_{1u},\tau_2,b_H}\varepsilon^{\sigma}-1\right)\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}\\
&+\gamma(1-\lambda_2)\,\beta\left(\frac{1}{2}\varepsilon^{\sigma}+\frac{1}{2}\mathbb{1}_{\varphi_{1u},\tau_0,b_L}(2-\varepsilon)^{\sigma}-1\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\\
&+\frac{\beta\gamma\lambda_2}{1-\beta}\left(\frac{\tau_2}{\tau_1}\right)^{1-\sigma}+\frac{\beta\gamma\lambda_2}{1-\beta}\beta\left(\frac{1}{2}\mathbb{1}_{\varphi_{1u},\tau_2,b_H}\varepsilon^{\sigma}-1\right)\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}
\end{aligned} \tag{25}
$$

Where $\max\{\pi(\varphi,\tau,b),0\}\equiv\mathbb{1}_{\varphi,\tau,b}\pi(\varphi,\tau,b)\equiv\tilde{\pi}(\varphi,\tau,b)$. There are three indicators $\mathbb{1}_{\varphi_{1u}\tau_1,b_H}$, $\mathbb{1}_{\varphi_{1u},\tau_2,b_H}$ and $\mathbb{1}_{\varphi_{1u},\tau_0,b_L}$ that are undetermined. In order to obtain an exact solution of $\frac{\varphi_{1u}}{\varphi_1}$, I need to discuss the value of three indicators case by case. In Appendix I, I show all 6 possible solutions. Based on the assumption such that entry cutoff firms φ_{1u} are willing to enter in low tariff state τ_0 but wait in high tariff state τ_2, the decision tree of φ_{1u} firms is illustrated in Figure 5. $\tilde{\pi}(\varphi_{1u},\tau_0,b_H)>0$ and φ_{1u} firms should be able to make positive profit in the best case scenario. Besides, $\tilde{\pi}(\varphi_{1u},\tau_2,b_L)=0$ and φ_{1u} firms cannot make positive profit in the worst case scenario.[41] If there is no TPU and $\tau_2=\tau_0=\tau_0$, φ_{1u} will be equal to φ_1.

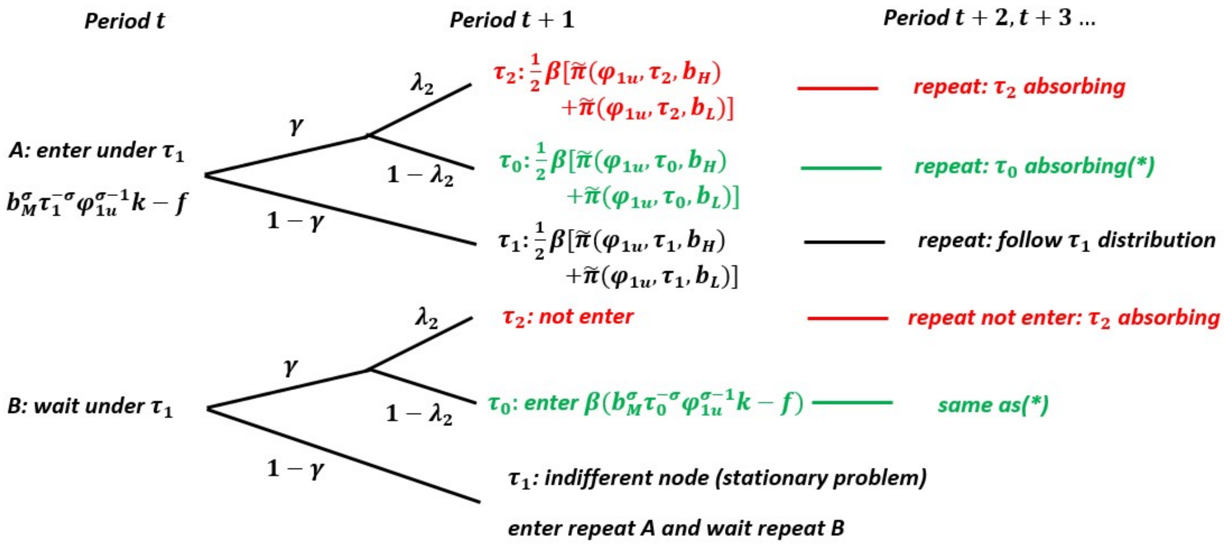

Fig. 5: Fixed cost decision tree

As before, in Figure 5, the red terms capture the effect of TPU through bad news τ_2 and the green terms capture the effect of TPU through good news τ_0. Increasing bad state tariff τ_2 weakly decreases the value of upper red terms while has no effect on the value of lower red terms. Therefore, bad news weakly deters firms' early entry. Reducing good state tariff τ_0 increases both the value of upper green terms and the value of lower green terms. However, since the relation between $\frac{1}{2}\left(b_H^{\sigma}+\mathbb{1}_{\varphi_{1u},\tau_0,b_L}b_L^{\sigma}\right)$ and b_M^{σ} can be ambiguous as $\mathbb{1}_{\varphi_{1u},\tau_0,b_L}=0$, the effect of good news is unclear. In order to better understand the relation between entry threshold under TPU - φ_{1u} and entry threshold under zero TPU - φ_1, I focus on the export decision of φ_1 firms under TPU. Recall that the expected value in

[41] If $\tilde{\pi}(\varphi_{1u},\tau_2,b_L)>0$, φ_{1u} firms will strictly prefer to enter under τ_1, which contradicts its definition.

intermediate tariff state τ_1 of a firm φ is

$$\Pi(\varphi, \tau_1) = \max\left\{\Pi_e(\varphi, \tau_1), \beta E_{\tau_1}\Pi(\varphi, \tau')\right\} \tag{26}$$

Where $\Pi_e(\varphi, \tau_1)$ is the value of entering under τ_1 and $\beta E_{\tau_1}\Pi(\varphi, \tau')$ is the value of waiting under τ_1. Define $Diff\,(\varphi, \tau_1) = \Pi_e(\varphi, \tau_1) - \beta E_{\tau_1}\Pi(\varphi, \tau')$ the net difference between the value of entering and the value of waiting under τ_1. If $Diff\,(\varphi, \tau_1) > 0$, firm φ will prefer to enter under τ_1. Otherwise firm φ will prefer to wait under τ_1. Therefore, as $Diff\,(\varphi_1, \tau_1) > 0$, φ_1 firms prefer to enter under TPU and there will be more entrants. As $Diff\,(\varphi_1, \tau_1) < 0$, φ_1 firms prefer to wait under TPU and there will be less entrants.[42] The sign of $Diff\,(\varphi_1, \tau_1)$ can be used to determine the relation between φ_1 and φ_{1u} while it gives less information on φ_{1u} itself. There is a following lemma.

Lemma 3

$Diff\,(\varphi_1, \tau_1)$ is proportional to $Sdiff\,(\varphi_1, \tau_1)$ which is

$$\begin{aligned} Sdiff\,(\varphi_1, \tau_1) = &- \frac{\pi\,(\varphi_1, \tau_1, b_H)}{1-\beta} + \frac{\lambda_2}{1-\beta}\tilde{\pi}\,(\varphi_1, \tau_2, b_H) \\ &+ (1-\lambda_2)\Big(\pi\,(\varphi_1, \tau_0, b_H) + \tilde{\pi}\,(\varphi_1, \tau_0, b_L) - 2\pi\,(\varphi_1, \tau_0, b_M)\Big) \end{aligned} \tag{27}$$

Proof: recall that φ_1 is entry threshold in no TPU case where trade policy is constant τ_1. Therefore, φ_1 firms are willing to enter in good state τ_0 and wait in bad state τ_2. I show that, assuming φ_1 firms are willing to enter(wait) under τ_1 as TPU is imposed, $Diff\,(\varphi_1, \tau_1)$ is proportional to $Sdiff\,(\varphi_1, \tau_1)$. For both assumptions - enter and wait under τ_1, the proportion is an increasing function of γ and β. See Appendix J for more details.

The probability of a tariff change γ doesn't affect the sign of $Diff\,(\varphi_1, \tau_1)$ and the intuition is the same as that in sunk cost learning model. $Sdiff\,(\varphi_1, \tau_1)$ can be rewritten as

$$Sdiff\,(\varphi_1, \tau_1) = L(\varphi_1, \tau_1) + B(\varphi_1, \tau_2) + G(\varphi_1, \tau_0) \tag{28}$$

Where

$$L(\varphi_1, \tau_1) = -\frac{\pi\,(\varphi_1, \tau_1, b_H)}{1-\beta} = -\frac{b_H^{\sigma}\tau_1^{-\sigma}\varphi_1^{\sigma-1}k - f}{1-\beta} \tag{29}$$

$$B(\varphi_1, \tau_2) = \frac{\lambda_2\tilde{\pi}\,(\varphi_1, \tau_2, b_H)}{1-\beta} = \frac{\lambda_2 \max\left\{b_H^{\sigma}\tau_2^{-\sigma}\varphi_1^{\sigma-1}k - f, 0\right\}}{1-\beta} \tag{30}$$

$$\begin{aligned} G(\varphi_1, \tau_0) &= (1-\lambda_2)\Big(\pi\,(\varphi_1, \tau_0, b_H) + \tilde{\pi}\,(\varphi_1, \tau_0, b_L) - 2\pi\,(\varphi_1, \tau_0, b_M)\Big) \\ &= (1-\lambda_2)\Big(b_H^{\sigma}\tau_0^{-\sigma}\varphi_1^{\sigma-1}k + \max\left\{b_L^{\sigma}\tau_0^{-\sigma}\varphi_1^{\sigma-1}k - f, 0\right\} - 2b_M^{\sigma}\tau_0^{-\sigma}\varphi_1^{\sigma-1}k + f\Big) \end{aligned} \tag{31}$$

$L(\varphi_1, \tau_1)$ captures the net loss in intermediate state τ_1. As TPU is introduced, φ_1 firms get access to τ_1 with less probability and earn less profits under τ_1 compared to the no TPU case. $B(\varphi_1, \tau_2)$

[42] In order to make $Diff\,(\varphi_1, \tau_1) > 0$ be equivalent to $\varphi_1 > \varphi_{1u}$ and $Diff\,(\varphi_1, \tau_1) < 0$ be equivalent to $\varphi_1 < \varphi_{1u}$, a condition such that $Diff\,(\varphi, \tau_1)$ passes through zero line once from below as φ increases should be imposed. A sufficient condition such that $Diff\,(\varphi, \tau_1)$ is an increasing function of φ is that $1 - \beta(1-\gamma) + \beta\gamma(1-\lambda_2)\left(\frac{1}{2}\varepsilon^{\sigma} - 1\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma} + \beta\gamma\lambda_2\left(\frac{1}{2}\varepsilon^{\sigma} - 1\right)\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma} > 0$.

and $G(\varphi_1, \tau_0)$ capture the net gain of bad news and good news separately.[43] As $\tau_2 = \tau_0 = \tau_1$, $Sdiff\left(\varphi_1, \tau_1\right) = 0$ and $\varphi_{1u} = \varphi_1$. $\frac{\partial B(\varphi_{1u}, \tau_2)}{\partial \tau_2}$ and $\frac{\partial G(\varphi_{1u}, \tau_0)}{\partial \tau_0}$ measure the marginal effect of TPU on φ_1 firms' relative profits of early entry through bad and good news separately. As in sunk cost learning model, bad news affects multiple periods while good news only affects one period. Therefore, there is a factor $\frac{1}{1-\beta}$ in bad news term $B(\varphi_1, \tau_2)$. There are 2 following lemmas.

Lemma 4

Bad news principal: for φ_1 firms, increasing bad state tariff τ_2 weakly decreases the relative profits of early entry.

Proof: from equation (30), it's easy to verify that $B(\varphi_1, \tau_2)$ is a weakly decreasing function of τ_2. Therefore, the net marginal effect of TPU through high tariff τ_2 on relative early entry profits is negative and bad news weakly deters φ_1 firms' entry.

Increasing bad state tariff τ_2 weakly reduces the net gain of bad news $B(\varphi_1, \tau_2)$ and early entry becomes less appealing for φ_1 firms. Using fixed cost f, the expected per period profit $\tilde{\pi}\left(\varphi_1, \tau_2, b_H\right)$ is bounded above 0. Therefore, as $\tilde{\pi}\left(\varphi_1, \tau_2, b_H\right) = 0$, bad news τ_2 has no more effect on φ_1 firms' profits and a higher τ_2 cannot deter φ_1 firms' early entry further.

Lemma 5

Good news principal: for φ_1 firms, decreasing good state tariff τ_0 may decrease the relative profits of early entry if good news τ_0 is not sufficiently low and good belief b_H is not sufficiently high. However, for a sufficiently low τ_0 or high b_H, decreasing τ_0 increases the relative profits of early entry.

Proof: from equation (31), $G(\varphi_1, \tau_0)$ is a decreasing function of τ_0 as $\tilde{\pi}\left(\varphi_1, \tau_0, b_L\right) > 0$ given that $\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma - b_M^\sigma > 0$. However, as $\tilde{\pi}\left(\varphi_1, \tau_0, b_L\right) = 0$ and $\frac{1}{2}b_H^\sigma - b_M^\sigma < 0$, $G(\varphi_1, \tau_0)$ is an increasing function of τ_0.

A counterintuitive result here is that good news can make early entry less appealing for φ_1 firms. As $\tilde{\pi}\left(\varphi_1, \tau_0, b_L\right) = 0$, conditional on entering under τ_1 in period t, in period $t+1$ under τ_0, φ_1 firms only make positive profits in the best case scenario where they hold good belief b_H, which happens with probability $\frac{1}{2}$. Considering the risk of having a bad belief b_L afterward and a relatively lower cost of entering under τ_0, if good belief b_H is not high enough, waiting in period t and entering in period $t+1$ under τ_0 can be a more appealing choice. In this case, a smaller τ_0 (better news) can favor waiting more. Recall that early entry makes firms benefit from the extra profit of learning one period earlier. As $\tilde{\pi}\left(\varphi_1, \tau_0, b_L\right) = 0$ and $\frac{1}{2}b_H^\sigma - b_M^\sigma < 0$, this extra benefit is not profitable enough and cannot compensate for the loss under τ_1. Since good news may deter φ_1 firms' early entry, for given τ_2 and τ_0, the effect of λ_2 can also be non monotone. Using a mean-preserving spread and substituting τ_2 by $\delta\tau_1$ and τ_0 by $\frac{1-\lambda_2\delta}{1-\lambda_2}\tau_1$, there is a following proposition and the intuition is similar to that of proposition 1.

Proposition 4

In fixed cost learning model, if TPU measure δ is small, $Sdiff\left(\varphi_1, \tau_1\right) < 0$ and there will be less entrants under TPU. However, if δ is large enough, $Sdiff\left(\varphi_1, \tau_1\right)$ can be greater than 0 and

[43] $L(\varphi_1, \tau_1)$ is not the real net loss and $B(\varphi_1, \tau_2)$ and $G(\varphi_1, \tau_0)$ are not the real net gain since $Sdiff\left(\varphi_1, \tau_1\right)$ is proportional to $Diff\left(\varphi_1, \tau_1\right)$.

there can be more entrants under TPU.

Proof: see Appendix K for more details.

Recall that in sunk cost learning model, as the variance measure of posterior beliefs $\varepsilon = 1$, there is no learning and $\frac{\varphi_{1u}}{\varphi_1}$ is H&L solution such that $\varphi_{1u} > \varphi_1$ and there are less entrants under TPU. However, using fixed cost, as $\varepsilon = 1$, firms' entry decision is just based on profit in the current period and there is no state dependence. In this case, $\varphi_{1u} = \varphi_1$ which is equal to zero profit cutoff of $\pi(\varphi, \tau_1, b_M)$. There is a following proposition.

Proposition 5

(1) As $\varepsilon = 1$, $Sdiff(\varphi_1, \tau_1) = 0$ and $\varphi_{1u} = \varphi_1$ which is equal to zero profit cutoff of $\pi(\varphi, \tau_1, b_M)$.
(2) As ε is close to 1, $\frac{\partial Sdiff(\varphi_1, \tau_1)}{\partial \varepsilon} < 0$ and $\varphi_{1u} > \varphi_1$. There are less entrants under TPU as there is not much demand learning.
(3) As $\lambda_2 \tau_2^{-\sigma} + (1-\lambda_2)\tau_0^{-\sigma} < \tau_1^{-\sigma}$, $\frac{\partial Sdiff(\varphi_1, \tau_1)}{\partial \varepsilon} < 0$ and there are always less entrants under TPU.
(4) As $\lambda_2 \tau_2^{-\sigma} + (1-\lambda_2)\tau_0^{-\sigma} > \tau_1^{-\sigma}$, $\frac{\partial Sdiff(\varphi_1, \tau_1)}{\partial \varepsilon} > 0$ if ε is large enough. In this case, there can be more entrants under TPU. If ε is not large enough, the sign of $\frac{\partial Sdiff(\varphi_1, \tau_1)}{\partial \varepsilon}$ is ambiguous.

Proof: see Appendix L.

Like in sunk cost learning model, if TPU process is unfavorable such that $\lambda_2 \tau_2^{-\sigma} + (1-\lambda_2)\tau_0^{-\sigma} < \tau_1^{-\sigma}$, there cannot be more entrants under TPU. In addition, using fixed cost, entry firms only care about the post-learning scenarios with positive expected profits. Therefore, the marginal effect of ε only acts on these positive scenarios. Increasing the variance measure of posterior beliefs ε can be relatively less favorable to early entry under TPU even if TPU process is favorable.

In this section, I use per period fixed cost instead of sunk entry cost. For both TPU and no TPU cases, entry cutoff firms make negative profit in the entry period and hope to receive a good demand signal afterward. Since per period export profit can be negative, there exist multiple kinks in the solution of $\frac{\varphi_{1u}}{\varphi_1}$ and the problem becomes less tractable. I use $Sdiff(\varphi_1, \tau_1)$ instead to study the relation between φ_{1u} and φ_1. Unlike sunk cost learning model, as exporters only take the scenarios with positive profit into account after learning, good news can deter firms' early entry in some cases. Some other results in sunk cost learning model are not robust in fixed cost learning model either.

5 Empirical application (very preliminary)

In this section, I will give a brief discussion on my empirical application. I have shown that in H&L partial equilibrium model, there is no good news effect and TPU only deters firms' entry through bad news τ_2. However, in my model, good news τ_0 matters. If I am able to find that good news matters empirically, H&L's result will be non-robust. More precisely, my sunk cost learning model predicts that the good news encourages entry.[44] I would like to test the effect of good news separately by shutting down bad news channel entirely. Empirically, it is difficult to distinguish between bad news effect and good news effect as they can happen at the same time. For example, if a Regional Trade Agreement (RTA) is enforced in the future between 2 countries, a potential exporter may expect that it is more (less) likely to face a lower (higher) tariff and the 2 effects jointly affect export decision. As I shut down bad news channel, using sunk cost learning model, my prediction is that for potential exporters that face 0 risk of future bad news (higher tariff), future good news (lower tariff) encourages their current entry in foreign market.

[44] Even though my fixed cost learning model predicts that the effect of good news on entry can be negative, I believe it happens less likely.

For WTO country pairs, bound tariffs can be nearly considered as the worst case scenario since it is the maximum MFN tariff level for a given commodity line.[45] Therefore, for product lines whose MFN tariff equals to bound tariff (BND), there is no effect of bad news as bound tariff is considered as the worst case scenario. I use the future enforcement of RTA agreement as a pure good shock to test good news effect for those products whose MFN=BND>0. Figure 6 and 7 illustrate decision tree as bad news channel is shut down ($\lambda_2 = 0$) using sunk cost learning model and H&L model separately. The green terms capture the effect through good news channel which is positive in sunk cost learning model and 0 in H&L model. Good news state τ_0 is the realization of RTA enforcement where $\tau_0 \leq \tau_1$.[46]

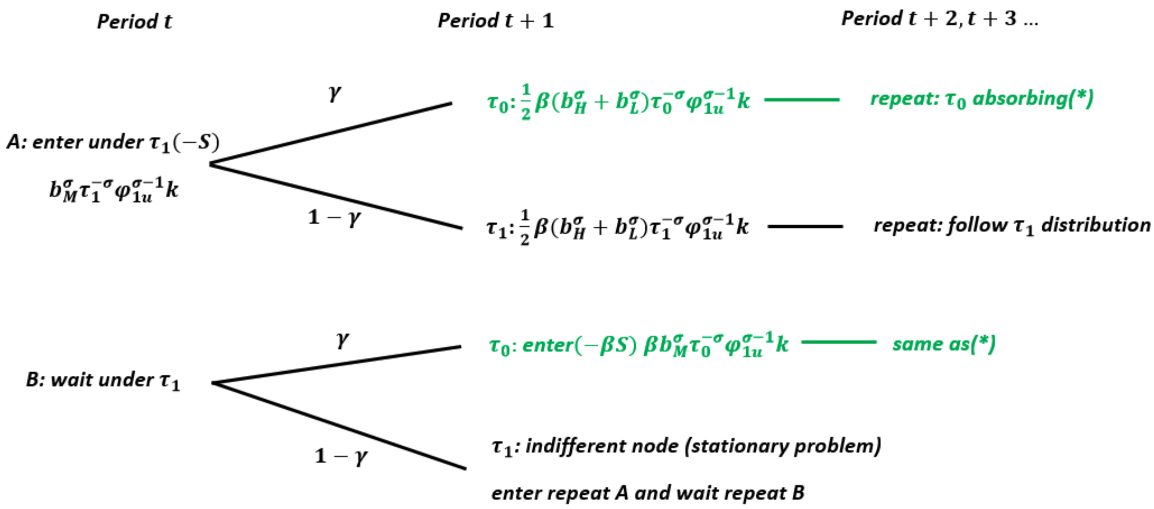

Fig. 6: Good news effect in sunk cost learning model

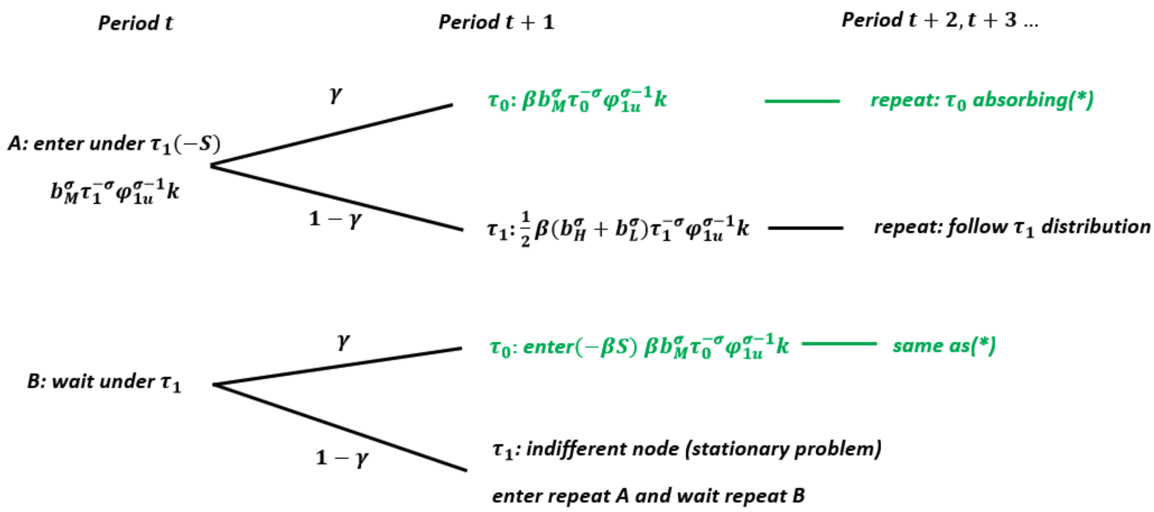

Fig. 7: Good news effect in H&L model

Two main assumptions are required. Firstly, if the current applied MFN tariff equals to bound tariff, there cannot be a worse news in the future ($\lambda_2 = 0$). So I exclude the possibility such that

[45] I say nearly because, in reality, applied tariff can be higher than bound tariff, e.g., temporary trade barriers.

[46] Recall that τ_0 is an absorbing state. Using the US and Canada as import country, in my dataset, there is no exit of RTA. Therefore, RTA enforcement can be approximately considered as an absorbing state.

applied tariff can be higher than bound tariff.[47] Secondly, as a RTA will be enforced in the future, it is more likely that future applied tariff decreases because of imposition of lower preferential tariff. In other words, as the enforcement date is closer, γ increases and a realization of good news state τ_0 becomes more likely.[48]

I use product entry as an indirect measure of firms' entry in foreign market and focus on two measures at different levels. The first is an indicator of positive HS6 trade flow which equals to 1 if there exists positive HS6 bilateral trade flow between 2 countries i and j at a given year t and 0 otherwise. The second is a product entry share at country-pair-year ijt level which is[49]

$$positive\_share_{ijt} = \frac{Number\,of\,products\,whose\,mfn = bnd\,with\,positive\,trade\,flow\,within\,ijt}{Number\,of\,products\,whose\,mfn = bnd\,within\,jt} \tag{32}$$

Using the first measure - product level entry, I cannot include very disaggregate dummy variables as there are too many observations. By aggregating product-level entry to country-pair-level entry share, I can include an exhaustive set of dummies and also perform two-way fixed effects estimation with heterogeneous treatment effects. Meanwhile, I lose entry information at product-level. The data sets I use are the BACI database, the Gravity database and the Trains database. The BACI database is used to build my two dependant variables - product level entry and country pair level entry share. The Gravity database provides information on WTO membership, RTA, etc. The Trains database gives information on MFN tariff and bound tariff at HS6 level. I only keep the US and Canada as destination since I believe these 2 countries have more complete trade flows records in the BACI database. The period of coverage is from year 1995 to year 2019 and there are 149 WTO export countries included.[50]

Before showing the regression results, let me mention some caveats. Firstly, I need to be careful about general equilibrium effect. Using a general equilibrium model, H&L point that if the current state is close to bad state (bad news is shut down), TPU can induce exporters' entry in foreign market. It is because good news to home exporters is bad news to foreign domestic firms. Foreign domestic firms exist, which pushes up foreign price and induces more exporters' entry in foreign market. Therefore, to control general equilibrium effect through aggregate price, adding importer-side dummy controls is necessary. Secondly, good news effect is not specific to learning model. Any model that creates age dependence is able to generate good news effect under TPU, e.g., demand accumulation. However, the underlying mechanism should be similar to what has been discussed in the above theoretical part. The key element to make good news matter is age dependence. Thirdly, applied tariff (AHS) may not be equal to MFN, e.g., temporary trade barrier for pre-RTA phase and preferential tariff for post-RTA phase. In the Trains HS6 bilateral tariff database, there is less information on applied tariffs. Using applied tariff, I cannot build an exhaustive database including all WTO countries for each product with MFN=BND and many 0 trade flows will not be included. Even if AHS information is not exhaustive, I can still use it as a robust check.[51] Fourthly, I should be careful about the endogeneity of RTA. Baier and Bergstrand (2007) state that introducing instrumental variables doesn't work well to deal with endogenous RTA problem. Using panel data,

[47] I also assume that bound tariff cannot increase. In the data, both MFN and bound tariff are quite stable over time.

[48] I currently focus on the enforcement date of RTA because I am able to get the information directly from the Gravity database. Baier and Bergstrand (2007) also use the "Date of Entry into Force" of the agreement to build the agreement dummy. The signature date of RTA also matters and affects γ, which will be studied in the future.

[49] Within destination-year jt, the number of products whose MFN=BND doesn't vary across export countries by definition of MFN.

[50] See Appendix M for more details on data construction.

[51] For example, if a bilateral applied tariff of some product is not documented in the Trains HS6 bilateral tariff database, I can assume that the applied tariff is just equal to MFN. Or I can build a smaller sample using applied tariff only.

introducing export country-year and destination-year fixed effect can deal with endogenous problem. Therefore, I should also include dummy controls for both export countries and destinations.

Below, I show three main regression results and other results are put in Appendix N. i is export country and j is destination. mfn_{jpt} is simple average MFN tariff at HS6 level and p denotes HS6 product. RTA_{ijt} is Regional Trade Agreement dummy. $RTA_{ijt} = 1$ if RTA is enforced at ijt level. $positive_{ijpt} = 1$ if positive trade flow at $ijpt$ level. $pre1_{ijt} = 1$ for one year before RTA enforcement and $pre1_{ijt} = 0$ for all the other years. $post1_{ijt} = 1$ for the first year of RTA enforcement and $post1_{ijt} = 0$ for all the other years. The same rule is applied to $pre2_{ijt} = 1$, $pre3_{ijt} = 1$, $post2_{ijt} = 1$, $post3_{ijt} = 0$. Gravity controls are $lnpop$, $lngdpcap$, $lndistw$, $Contiguity$, $common\,language$ and $common\,legal\,origins\,before\,/\,after\,transition$.[52] it is export country-year dummy. jt is destination-year dummy. ij is export country-destination dummy. $jhs2t$ is destination-HS2 sector-year dummy.

5.1 Product-level entry

$$\begin{aligned} D.positive_{ijpt} = &\beta_1 * D.mfn_{jpt} + \beta_2 * D.pre1_{ijt} + \beta_3 * D.pre2_{ijt} + \beta_4 * D.pre3_{ijt} + \beta_5 * D.post1_{ijt} \\ &+ \beta_6 * D.post2_{ijt} + \beta_7 * D.post3_{ijt} + \beta_g * D.gravity + \beta_d * Dummy + \epsilon_{ijpt} \end{aligned} \tag{33}$$

Table 3 illustrates the result of first-difference regression which captures dynamic effect for 3 periods pre and post RTA. $D.mfn_{jpt}$ is 0 because at product level, MFN tariff doesn't vary across years. The positive one period pre-RTA effect is robust across different specifications, which suggests that good news effect can be positive. Surprisingly, the post-RTA effect is non significant across different specifications.[53]

[52] See Appendix M for the description of gravity controls. For EU population, I take the sum. For EU gdp per capita, weighted distance, contiguity, common language, common legal origins and RTA, I just take simple average.

[53] It doesn't mean that post-RTA has no effect on trade. Here I only consider extensive margin - product entry and RTA can also have an effect on intensive margin.

Tab. 3: Dynamics effect of RTA - 3 periods before and 3 periods after first difference

	(1) D.positive	(2) D.positive	(3) D.positive	(4) D.positive
D.mfn	0 (.)	0 (.)	0 (.)	0 (.)
D.pre1	0.00795** (2.24)	0.00754** (2.17)	0.0105* (1.73)	0.0105* (1.73)
D.pre2	0.000677 (0.22)	0.000416 (0.13)	-0.00632* (-1.74)	-0.00632* (-1.74)
D.pre3	0.00174 (0.52)	0.00169 (0.52)	-0.00148 (-0.41)	-0.00148 (-0.41)
D.post1	0.00440 (1.09)	0.00407 (1.04)	-0.00252 (-0.56)	-0.00252 (-0.56)
D.post2	0.00599 (1.41)	0.00553 (1.34)	0.00256 (0.56)	0.00256 (0.56)
D.post3	-0.000960 (-0.11)	-0.00120 (-0.14)	-0.00359 (-0.53)	-0.00359 (-0.53)
D.lnpop_o		0.00192 (0.10)		
D.lnpop_d		0.00550 (0.29)		
D.lngdpcap_o		-0.00602** (-2.00)		
D.lngdpcap_d		0.0283*** (4.94)		
fixed effect			it, jt	it, jhs2t
cluster	ij	ij	ij	ij
Observations	5040518	5000542	5040518	5040518

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.2 Country-level entry share

$$\begin{aligned} D.positive\_share_{ijt} =& \beta_1 * D.m\_mfn_{jt} + \beta_2 * D.pre1_{ijt} + \beta_3 * D.pre2_{ijt} + \beta_4 * D.pre3_{ijt} + \beta_5 * D.post1_{ijt} \\ &+ \beta_6 * D.post2_{ijt} + \beta_7 * D.post3_{ijt} + \beta_g * D.gravity_{it/jt} + \beta_d * Dummy + \epsilon_{ijt} \end{aligned} \tag{34}$$

Table 4 also illustrates the result of first difference regression which captures dynamic effect for 3 periods pre and post RTA. m_mfn_{jt} is simple average of MFN of the products whose MFN=BND within jt. $D.m\_mfn_{jt}$ is not 0 because at destination level j, the set of products with MFN=BND changes across years. Using it and jt dummy controls, the effect of MFN tariff becomes negative. The positive one period pre-RTA effect is still robust across different specifications. And the post-RTA effect is non significant across different specifications.

Tab. 4: Dynamics effect of RTA - 3 periods before and 3 periods after first difference

	(1) D.positive_share	(2) D.positive_share	(3) D.positive_share
D.m_mfn	0.00384***	0.00265***	-0.297***
	(4.27)	(3.14)	(-2.75)
D.pre1	0.0150***	0.0142***	0.0139*
	(3.07)	(2.84)	(1.92)
D.pre2	0.00448	0.00494	-0.0106**
	(1.12)	(1.27)	(-2.28)
D.pre3	0.00326	0.00401	0.00270
	(1.00)	(1.16)	(0.66)
D.post1	0.00522	0.00563	-0.00423
	(0.83)	(0.92)	(-0.68)
D.post2	0.00963	0.00931	0.000428
	(1.51)	(1.50)	(0.06)
D.post3	0.00281	0.00231	-0.00621
	(0.38)	(0.31)	(-0.64)
D.lnpop_o		-0.0293	
		(-1.49)	
D.lnpop_d		0.0181	
		(0.79)	
D.lngdpcap_o		-0.0103***	
		(-3.34)	
D.lngdpcap_d		0.0851***	
		(9.84)	
fixed effect			it, jt
cluster	ij	ij	ij
Observations	5181	5133	5181

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Below I use two-way fixed effects estimation with heterogeneous treatment effects.[54] The control variables I include are m_mfn, gravity controls of export countries and destination-year dummy jt.[55] The positive one period and three period pre-RTA effects are significant at 95% level. And the post-RTA dynamic effect is non significant. Combining all results above, the reaction of product entry is positive even before the enforcement of RTA, which suggests that good news matters.

[54] The lth placebo compares first-time switchers' and not-yet switchers' outcome evolution, from the l+1th to the lth period before first-time switchers' treatment changes. dynamic() gives the number of dynamic treatment effects to be estimated. CI is 95% confidence interval.

[55] I don't include export country-year dummy *it* because it creates more than 3000 dummies. I tried to add them and didn't get the result after waiting for 10 days.

```
. did_multiplegt positive_share ij t rta , robust_dynamic dynamic(2)  placebo(
> 3)  jointtestplacebo controls(m_mfn lnpop_o lngdpcap_o jt2-jt50)  breps(2000
> ) cluster(ij) covariances save_results(W:\empirical test\baci_test_92\US_CAN
> _importer\2way_share_mfn_gravity_jt_USCAN.dta)

DID estimators of the instantaneous treatment effect, of dynamic treatment
effects if the dynamic option is used, and of placebo tests of the parallel
trends assumption if the placebo option is used. The estimators are robust
to heterogeneous effects, and to dynamic effects if the robust_dynamic
option is used.
```

	Estimate	SE	LB CI	UB CI	N
Effect_0	-.0046276	.003908	-.0122873	.0030322	3418
Effect_1	.0012554	.0048466	-.008244	.0107548	3150
Effect_2	-.0011486	.0083959	-.0176046	.0153074	2867
Placebo_1	.0126449	.0047842	.0032678	.0220219	3096
Placebo_2	.002765	.0040599	-.0051923	.0107223	2858
Placebo_3	.0071738	.0034328	.0004454	.0139022	2822

	Switchers
Effect_0	39
Effect_1	38
Effect_2	32
Placebo_1	38
Placebo_2	36
Placebo_3	36

6 Conclusion

In this paper, I introduce one-period demand learning in partial equilibrium version of Handley and Limão (2017) model. Their model gives a closed-form solution of entry cutoff and predicts that TPU reduces exporters' entry due to *bad news principle*. I study the effect of TPU on entry cutoff in the presence of demand learning. Two models have been studied, namely sunk cost learning model and fixed cost learning model. Like Handley and Limão (2017), bad news discourages firms' early entry. Moreover, good news matters in my model. My sunk cost learning model nests Handley and Limão model and is also able to give a closed-form solution of entry cutoff. In sunk cost learning model, good news encourages firms' early entry. Future possible good news favors early entry more than late entry as good news accompanied by early learning brings extra profit to early entry. In fixed cost learning model, per period export profit can be negative and there exists endogenous exit. The model becomes less tractable because of existence of multiple kinks. Moreover, good news can deter firms' early entry as entry cutoff firms take only the best learning outcome into account. I also study the effect of pure uncertainty using a mean-preserving spread in tariff. For a sufficiently large uncertainty, there can be more entrants under TPU. In my models, firms' willingness of making early entry under TPU jointly depends on how profitable the learning is, namely the variance of posterior beliefs, and how favorable the TPU process is, namely the weighted inverted sum of high and low tariffs net of intermediate tariff. For a non-favorable TPU process, even though demand learning can bring extra profit to early entry, there are relatively less entrants under TPU compared to no TPU case. The variance of posterior beliefs is negatively correlated with firms' exporting ages. For new exporters, they learn more from exporting than old ones do and their variance of posterior beliefs is larger. Therefore, my model suggests that the effect of TPU on exporters' entry decision is heterogeneous across exporters' age. As both literature of export under TPU and literature of firm dynamics are recent, my paper combines the 2 literature and is able to provide some new insights on export decision under TPU. In addition, my model can also be used to study other irreversible investment problems.

There exist some limitations in my model. I treat TPU in a very simple way and my model cannot be used to study other kinds of trade barrier such as quota and sanitary measures. Besides, trade policy uncertainty is assumed to be an exogenous process while tariff change can be an endogenous decision made by the government. I also assume that there is no correlation between learning process and TPU process. I leave for future research the post entry dynamics in the presence of TPU and multiple-periods demand learning. I also suggest to study entry decision under TPU using other export dynamics mechanisms.

References

Aeberhardt, Romain, Ines Buono, and Harald Fadinger (2014). "Learning, incomplete contracts and export dynamics: Theory and evidence from French firms". In: *European Economic Review* 68, pp. 219–249.

Albornoz, Facundo et al. (2012). "Sequential exporting". In: *Journal of International Economics* 88.1, pp. 17–31.

Alessandria, George A, Costas Arkolakis, and Kim J Ruhl (2020). *Firm dynamics and trade*. Tech. rep. National Bureau of Economic Research.

Alessandria, George A, Shafaat Y Khan, and Armen Khederlarian (2019). *Taking Stock of Trade Policy Uncertainty: Evidence from China's Pre-WTO Accession*. Tech. rep. National Bureau of Economic Research.

Arkolakis, Costas, Theodore Papageorgiou, and Olga A Timoshenko (2018). "Firm learning and growth". In: *Review of Economic Dynamics* 27, pp. 146–168.

Bacchetta, Marc and Roberta Piermartini (2011). *The value of bindings*. Tech. rep. WTO staff working paper.

Baker, Scott R, Nicholas Bloom, and Steven J Davis (2016). "Measuring economic policy uncertainty". In: *The quarterly journal of economics* 131.4, pp. 1593–1636.

Baley, Isaac, Laura Veldkamp, and Michael Waugh (2020). "Can global uncertainty promote international trade?" In: *Journal of International Economics* 126, p. 103347.

Bastos, Paulo, Daniel A Dias, and Olga A Timoshenko (2018). "Learning, prices and firm dynamics". In: *Canadian Journal of Economics/Revue canadienne d'économique* 51.4, pp. 1257–1311.

Berman, Nicolas, Vincent Rebeyrol, and Vincent Vicard (2019). "Demand learning and firm dynamics: evidence from exporters". In: *Review of Economics and Statistics* 101.1, pp. 91–106.

Bernanke, Ben S (1983). "Irreversibility, uncertainty, and cyclical investment". In: *The quarterly journal of economics* 98.1, pp. 85–106.

Caldara, Dario et al. (2020). "The economic effects of trade policy uncertainty". In: *Journal of Monetary Economics* 109, pp. 38–59.

Chen, Cheng et al. (2019). *Uncertainty, Imperfect Information and Learning in the International Market*. Tech. rep. working paper.

Crowley, Meredith, Ning Meng, and Huasheng Song (2018). "Tariff scares: Trade policy uncertainty and foreign market entry by Chinese firms". In: *Journal of International Economics* 114, pp. 96–115.

Das, Sanghamitra, Mark J Roberts, and James R Tybout (2007). "Market entry costs, producer heterogeneity, and export dynamics". In: *Econometrica* 75.3, pp. 837–873.

Eaton, Jonathan, Marcela Eslava, Cornell J Krizan, et al. (2014). "A search and learning model of export dynamics". In: *Unpublished manuscript*.

Eaton, Jonathan, Marcela Eslava, Maurice Kugler, et al. (2007). *Export dynamics in Colombia: Firm-level evidence*. Tech. rep. National Bureau of Economic Research.

Feng, Ling, Zhiyuan Li, and Deborah L Swenson (2017). "Trade policy uncertainty and exports: Evidence from China's WTO accession". In: *Journal of International Economics* 106, pp. 20–36.

Fitzgerald, Doireann, Stefanie Haller, and Yaniv Yedid-Levi (2019). *How exporters grow*. Tech. rep. working paper.

Foster, Lucia, John Haltiwanger, and Chad Syverson (2016). "The slow growth of new plants: learning about demand?" In: *Economica* 83.329, pp. 91–129.

Handley, Kyle (2014). "Exporting under trade policy uncertainty: Theory and evidence". In: *Journal of international Economics* 94.1, pp. 50–66.

Handley, Kyle and Nuno Limão (Nov. 2015). "Trade and Investment under Policy Uncertainty: Theory and Firm Evidence". In: *American Economic Journal: Economic Policy* 7.4, pp. 189–222.

Handley, Kyle and Nuno Limão (Sept. 2017). "Policy Uncertainty, Trade, and Welfare: Theory and Evidence for China and the United States". In: *American Economic Review* 107.9, pp. 2731–83.

Jovanovic, Boyan (1982). "Selection and the Evolution of Industry". In: *Econometrica: Journal of the Econometric Society*, pp. 649–670.

Melitz, Marc J (2003). "The impact of trade on intra-industry reallocations and aggregate industry productivity". In: *econometrica* 71.6, pp. 1695–1725.

Piveteau, Paul (2016). "An empirical dynamic model of trade with consumer accumulation". In: *American Economic Journal: Microeconomics*.

Ruhl, Kim J and Jonathan L Willis (2017). "New exporter dynamics". In: *International Economic Review* 58.3, pp. 703–726.

Steinberg, Joseph B (2019). "Brexit and the macroeconomic impact of trade policy uncertainty". In: *Journal of International Economics* 117, pp. 175–195.

Timoshenko, Olga A (Oct. 2015a). "Learning versus sunk costs explanations of export persistence". In: *European Economic Review* 79, pp. 113–128.

Timoshenko, Olga A (Mar. 2015b). "Product switching in a model of learning". In: *Journal of International Economics* 95.2, pp. 233–249.

Appendix

A Calculate the belief

(i) For a firm that starts to export in period t for the first time, the unconditional probability of her demand shock realization a_t will be:

$$Pr(a_t = \theta_H) = Pr(\theta = \theta_H, a_t = \theta_H) + Pr(\theta = \theta_L, a_t = \theta_H) = \frac{1}{2} \cdot p + \frac{1}{2} \cdot (1-p) = \frac{1}{2} \tag{1}$$

$$Pr(a_t = \theta_L) = Pr(\theta = \theta_H, a_t = \theta_L) + Pr(\theta = \theta_L, a_t = \theta_L) = \frac{1}{2} \cdot (1-p) + \frac{1}{2} \cdot p = \frac{1}{2} \tag{2}$$

Therefore, her belief (prior) of demand shock realization a_t in period t will be:

$$b_t = Pr(a_t = \theta_H) e^{\frac{\theta_H}{\sigma}} + Pr(a_t = \theta_L) e^{\frac{\theta_L}{\sigma}} = \frac{1}{2} e^{\frac{\theta_H}{\sigma}} + \frac{1}{2} e^{\frac{\theta_L}{\sigma}} = b_M, \tag{3}$$

(ii) For a firm that has exported in period $t-i$ with $i \geq 1$ for the first time and observed a realized demand shock $\bar{a}_t = a_{t-i} = \theta_H$, the conditional probability of her true underlying demand parameter θ will be

$$Pr(\theta = \theta_H | \bar{a}_t = \theta_H) = \frac{Pr(\bar{a}_t = \theta_H | \theta = \theta_H) \cdot Pr(\theta = \theta_H)}{Pr(\bar{a}_t = \theta_H, \theta = \theta_H) + Pr(\bar{a}_t = \theta_H, \theta = \theta_L)} = \frac{p \cdot \frac{1}{2}}{p \cdot \frac{1}{2} + (1-p) \cdot \frac{1}{2}} = p \tag{4}$$

$$Pr(\theta = \theta_L | \bar{a}_t = \theta_H) = \frac{Pr(\bar{a}_t = \theta_H | \theta = \theta_L) \cdot Pr(\theta = \theta_L)}{Pr(\bar{a}_t = \theta_H, \theta = \theta_H) + Pr(\bar{a}_t = \theta_H, \theta = \theta_L)} = \frac{(1-p) \cdot \frac{1}{2}}{p \cdot \frac{1}{2} + (1-p) \cdot \frac{1}{2}} = 1-p \tag{5}$$

Therefore, her posterior belief of a_t in period t will be

$$\begin{aligned} b_t(\bar{a}_t = \theta_H) &= Pr(\theta = \theta_H | \bar{a}_t = \theta_H) e^{\frac{\theta_H}{\sigma}} + Pr(\theta = \theta_L | \bar{a}_t = \theta_H) e^{\frac{\theta_L}{\sigma}} \\ &= p \cdot e^{\frac{\theta_H}{\sigma}} + (1-p) \cdot e^{\frac{\theta_L}{\sigma}} = b_H, \end{aligned} \tag{6}$$

(iii) For a firm that has exported in period $t-i$ with $i \geq 1$ for the first time and observed a realized demand shock $\bar{a}_t = a_{t-i} = \theta_L$, the conditional probability of her true underlying demand parameter θ will be

$$Pr(\theta = \theta_H | \bar{a}_t = \theta_L) = \frac{Pr(\bar{a}_t = \theta_L | \theta = \theta_H) \cdot Pr(\theta = \theta_H)}{Pr(\bar{a}_t = \theta_L, \theta = \theta_H) + Pr(\bar{a}_t = \theta_L, \theta = \theta_L)} = \frac{(1-p) \cdot \frac{1}{2}}{(1-p) \cdot \frac{1}{2} + p \cdot \frac{1}{2}} = 1-p \tag{7}$$

$$Pr(\theta = \theta_L | \bar{a}_t = \theta_L) = \frac{Pr(\bar{a}_t = \theta_L | \theta = \theta_L) \cdot Pr(\theta = \theta_L)}{Pr(\bar{a}_t = \theta_L, \theta = \theta_H) + Pr(\bar{a}_t = \theta_L, \theta = \theta_L)} = \frac{p \cdot \frac{1}{2}}{(1-p) \cdot \frac{1}{2} + p \cdot \frac{1}{2}} = p \tag{8}$$

Therefore, her posterior belief of a_t in period t will be

$$\begin{aligned} b_t(\bar{a}_t = \theta_L) &= Pr(\theta = \theta_H | \bar{a}_t = \theta_L) e^{\frac{\theta_H}{\sigma}} + Pr(\theta = \theta_L | \bar{a}_t = \theta_L) e^{\frac{\theta_L}{\sigma}} \\ &= (1-p) \cdot e^{\frac{\theta_H}{\sigma}} + p \cdot e^{\frac{\theta_L}{\sigma}} = b_L, \end{aligned} \tag{9}$$

Since p is assumed to be greater than $\frac{1}{2}$, $b_H \geq b_M \geq b_L > 0$ and $b_H + b_L = 2b_M$[1]. b_H increases in p and b_L decreases in p. As $p = \frac{1}{2}$, $b_H = b_L = b_M$ and exporters learn nothing from their export

[1] Using Jovanovic (1982) learning process, belief series is a martingale. Using our one period learning process, we can still capture the main idea of learning. Meanwhile, we are able to simplify our model as much as possible.

experience in the entry period. In this case, potential entrant is more like a very experienced exporter since receiving a new signal has no impact on her belief. As p increases, b_H and b_L diverge further from b_M and learning brings more information about exporters' type. In this case, potential entrant is more like a less experienced exporter since receiving a new signal has a large impact on her belief. The variance of belief sheds light on the age of exporter. However, we should keep in mind that more experienced incumbents are also highly selected. Therefore, their belief and productivity should be generally higher than that of fresh exporters. As $p = 1$ and $\theta_H \gg \theta_L$, b_H is close to $2b_M$, which also means that b_H cannot be greater than $2b_M$ by construction.

B Recursive form of $\Pi_e(\varphi, \tau)$

In this section, I try to rewrite $\Pi_e(\varphi, \tau)$ using recursive formula like H&L. The expected value from exporting after entry under τ_1 is

$$\begin{aligned}\Pi_e(\varphi, \tau_1) = b_M^\sigma \tau_1^{-\sigma}\Phi &+ \gamma\lambda_2 \frac{\beta}{1-\beta}\left(\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma\right)\tau_2^{-\sigma}\Phi \\ &+ \gamma\left(1-\lambda_2\right)\frac{\beta}{1-\beta}\left(\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma\right)\tau_0^{-\sigma}\Phi \\ &+ \left(1-\gamma\right)\beta\left(\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma\right)\tau_1^{-\sigma}\Phi + \left(1-\gamma\right)\beta\Re\left(\varphi, \tau_1\right) \\ = b_M^\sigma \tau_1^{-\sigma}\Phi &+ \frac{\gamma\lambda_2}{1-\left(1-\gamma\right)\beta}\frac{\beta}{1-\beta}\left(\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma\right)\tau_2^{-\sigma}\Phi \\ &+ \frac{\gamma\left(1-\lambda_2\right)}{1-\left(1-\gamma\right)\beta}\frac{\beta}{1-\beta}\left(\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma\right)\tau_0^{-\sigma}\Phi \\ &+ \frac{\left(1-\gamma\right)\beta}{1-\left(1-\gamma\right)\beta}\left(\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma\right)\tau_1^{-\sigma}\Phi\end{aligned} \tag{10}$$

Where $\Phi = \varphi^{\sigma-1}k$. Firm φ enters under τ_1 with unconditional belief b_M. With probability $\gamma\lambda_2$, trade policy realization will be τ_2 in period 2 and it will be τ_2 forever. We have the term $\gamma\lambda_2\frac{\beta}{1-\beta}\left(\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma\right)\tau_2^{-\sigma}\Phi$. With probability $\gamma\left(1-\lambda_2\right)$, trade policy realization will be τ_0 in period 2 and it will be τ_0 forever. We have the term $\gamma\left(1-\lambda_2\right)\frac{\beta}{1-\beta}\left(\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma\right)\tau_0^{-\sigma}\Phi$. With probability $1-\gamma$, trade policy realization will be τ_1 in period 2 and we have the term $\left(1-\gamma\right)\beta\left(\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma\right)\tau_1^{-\sigma}\Phi$. $\Re\left(\varphi, \tau_1\right)$ is the recursive term which means if trade realization in period 2 is τ_1, then in period 3, it will repeat the same process as that in period 2.[2] We have

$$\begin{aligned}\Re\left(\varphi, \tau_1\right) =&\gamma\lambda_2 \frac{\beta}{1-\beta}\left(\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma\right)\tau_2^{-\sigma}\Phi \\ &+ \gamma\left(1-\lambda_2\right)\frac{\beta}{1-\beta}\left(\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma\right)\tau_0^{-\sigma}\Phi \\ &+ \left(1-\gamma\right)\beta\left(\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma\right)\tau_1^{-\sigma}\Phi + \left(1-\gamma\right)\beta\Re\left(\varphi, \tau_1\right)\end{aligned} \tag{11}$$

If the entry condition is τ_2, we will have

$$\Pi_e(\varphi, \tau_2) = b_M^\sigma \tau_2^{-\sigma}\Phi + \frac{\beta}{1-\beta}\left(\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma\right)\tau_2^{-\sigma}\Phi \tag{12}$$

[2] It is because learning only lasts one period in my model.

Firm φ enters under τ_2 with unconditional belief b_M. As trade policy realization will be τ_2 forever, there is a term $\frac{\beta}{1-\beta}\left(\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma\right)\tau_2^{-\sigma}\Phi$. If the entry condition is τ_0, we will have

$$\Pi_e(\varphi, \tau_0) = b_M^\sigma \tau_0^{-\sigma}\Phi + \frac{\beta}{1-\beta}\left(\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma\right)\tau_0^{-\sigma}\Phi \tag{13}$$

Firm φ enters under τ_0 with unconditional belief b_M. Trade policy realization will be τ_0 forever. Then we can write $\Pi_e(\varphi, \tau_1)$, $\Pi_e(\varphi, \tau_2)$, $\Pi_e(\varphi, \tau_0)$ recursively.

$$\begin{aligned}\Pi_e(\varphi, \tau_1) =& b_M^\sigma \tau_1^{-\sigma}\Phi + (1-\gamma)\,\beta\left(\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma - b_M^\sigma\right)\tau_1^{-\sigma}\Phi \\ &+ \gamma\lambda_2\beta\left(\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma - b_M^\sigma\right)\tau_2^{-\sigma}\Phi \\ &+ \gamma\,(1-\lambda_2)\,\beta\left(\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma - b_M^\sigma\right)\tau_0^{-\sigma}\Phi \\ &+ \beta E_{\tau_1}\Pi_e(\varphi, \tau') \\ =& f(\varphi, \tau_1) + \beta E_{\tau_1}\Pi_e(\varphi, \tau')\end{aligned} \tag{14}$$

Where $f(\varphi, \tau_1)$ is

$$\begin{aligned}f(\varphi, \tau_1) =& b_M^\sigma \tau_1^{-\sigma}\Phi + (1-\gamma)\,\beta\left(\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma - b_M^\sigma\right)\tau_1^{-\sigma}\Phi \\ &+ \gamma\lambda_2\beta\left(\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma - b_M^\sigma\right)\tau_2^{-\sigma}\Phi \\ &+ \gamma\,(1-\lambda_2)\,\beta\left(\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma - b_M^\sigma\right)\tau_0^{-\sigma}\Phi\end{aligned} \tag{15}$$

$$\begin{aligned}\Pi_e(\varphi, \tau_2) =& b_M^\sigma \tau_2^{-\sigma}\Phi + \beta\left(\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma - b_M^\sigma\right)\tau_2^{-\sigma}\Phi + \beta E_{\tau_2}\Pi_e(\varphi, \tau') \\ =& f(\varphi, \tau_2) + \beta E_{\tau_2}\Pi_e(\varphi, \tau') \\ =& f(\varphi, \tau_2) + \beta\Pi_e(\varphi, \tau_2)\end{aligned} \tag{16}$$

Where $f(\varphi, \tau_2)$ is

$$f(\varphi, \tau_2) = b_M^\sigma \tau_2^{-\sigma}\Phi + \beta\left(\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma - b_M^\sigma\right)\tau_2^{-\sigma}\Phi \tag{17}$$

$$\begin{aligned}\Pi_e(\varphi, \tau_0) =& b_M^\sigma \tau_0^{-\sigma}\Phi + \beta\left(\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma - b_M^\sigma\right)\tau_0^{-\sigma}\Phi + \beta E_{\tau_0}\Pi_e(\varphi, \tau') \\ =& f(\varphi, \tau_0) + \beta E_{\tau_0}\Pi_e(\varphi, \tau') \\ =& f(\varphi, \tau_0) + \beta\Pi_e(\varphi, \tau_0)\end{aligned} \tag{18}$$

Where $f(\varphi, \tau_0)$ is

$$f(\varphi, \tau_0) = b_M^\sigma \tau_0^{-\sigma}\Phi + \beta\left(\frac{1}{2}b_H^\sigma + \frac{1}{2}b_L^\sigma - b_M^\sigma\right)\tau_0^{-\sigma}\Phi \tag{19}$$

The general expression of $\Pi_e(\varphi, \tau)$ is

$$\Pi_e(\varphi, \tau) = f(\varphi, \tau) + \beta E_\tau \Pi_e(\varphi, \tau') \tag{20}$$

$f(\varphi, \tau)$ captures the difference between the expected value from exporting after entry in the current period $\Pi_e(\varphi, \tau)$ and the expected value from exporting after entry in the next period $\beta E_\tau \Pi_e(\varphi, \tau')$.

C The monotonicity of $V(\varphi, \tau)$

Ideally, we want $V(\varphi, \tau)$ to be a decreasing function of φ and increasing function of τ.

(A) For any given φ, $V(\varphi, \tau)$ is an increasing function of τ.

(i) $f(\varphi, \tau)$ is a decreasing function of τ. We can easily observe that $f(\varphi, \tau_0) > f(\varphi, \tau_1) > f(\varphi, \tau_2)$. Therefore, $-f(\varphi, \tau)$ is an increasing function of τ.

(ii) Because of the first-order dominance, $E_{\tau_0} V(\varphi, \tau') < E_{\tau_1} V(\varphi, \tau') < E_{\tau_2} V(\varphi, \tau')$ if $V(\varphi, \tau)$ is increasing in τ.

By property (i) and (ii), if we start with an increasing $V(\varphi, \tau)$ in τ, the fixed point to this iteration is also increasing τ.[3]

(B) For any given τ, $V(\varphi, \tau)$ is a decreasing function of φ.

In H&L (2017), they don't give such proof. But the logic should be like this: $f(\varphi, \tau)$ is an increasing function of φ, which means $-f(\varphi, \tau)$ is a decreasing function of φ. Then starting with a decreasing $V(\varphi, \tau)$ in φ, the fixed point to this iteration is also decreasing in φ.

Therefore, firm will be more willing to choose to wait in the current period if their productivity φ is low and the current trade policy realization τ is high.

D Calculate $\frac{\varphi_{1u}}{\varphi_1}$ using sunk cost learning model

φ_{1u} is the entry threshold under TPU which satisfies the following equation

$$\beta E_{\tau_1} V(\varphi_{1u}, \tau') - f(\varphi_{1u}, \tau_1) + (1 - \beta) S = 0 \tag{21}$$

Recall that $V(\varphi_{1u}, \tau_0) = 0$, $V(\varphi_{1u}, \tau_1) = 0$ and $V(\varphi_{1u}, \tau_2) > 0$. $E_{\tau_1} V(\varphi_{1u}, \tau')$ will be

$$\begin{aligned} E_{\tau_1} V(\varphi_{1u}, \tau') &= \gamma \lambda_2 V(\varphi_{1u}, \tau_2) \\ &= \gamma \lambda_2 \left(\beta E_{\tau_2} V(\varphi_{1u}, \tau') - f(\varphi_{1u}, \tau_2) + (1 - \beta) S \right) \end{aligned} \tag{22}$$

We have

$$\begin{aligned} E_{\tau_2} V(\varphi_{1u}, \tau') &= V(\varphi_{1u}, \tau_2) \\ &= \beta E_{\tau_2} V(\varphi_{1u}, \tau') - f(\varphi_{1u}, \tau_2) + (1 - \beta) S \end{aligned} \tag{23}$$

From (23), we can solve $E_{\tau_2} V(\varphi_{1u}, \tau')$

$$E_{\tau_2} V(\varphi_{1u}, \tau') = S - \frac{f(\varphi_{1u}, \tau_2)}{1 - \beta} \tag{24}$$

Bring (24) into (22) and we have

$$E_{\tau_1} V(\varphi_{1u}, \tau') = \gamma \lambda_2 \left(S - \frac{f(\varphi_{1u}, \tau_2)}{1 - \beta} \right) \tag{25}$$

[3] A difference between my model and that of H&L (2017) is that their $f(\varphi, \tau)$ is the per period profit which has the same form for different τ.

Bring (25) into (21) and we have

$$\beta\gamma\lambda_2\left(S-\frac{f(\varphi_{1u},\tau_2)}{1-\beta}\right)-f(\varphi_{1u},\tau_1)+(1-\beta)\,S=0 \tag{26}$$

Arrange it as

$$f(\varphi_{1u},\tau_1)+\frac{\beta\gamma\lambda_2}{1-\beta}f(\varphi_{1u},\tau_2)=(1-\beta+\beta\gamma\lambda_2)\,S \tag{27}$$

Recall that φ_1 is the entry threshold in a no TPU case and we have

$$b_M^\sigma\tau_1^{1-\sigma}\Phi_1+\frac{1}{2}\frac{\beta}{1-\beta}\left(b_H^\sigma+b_L^\sigma\right)\tau_1^{1-\sigma}\Phi_1=S \tag{28}$$

Where $\Phi_1=\varphi_1^{\sigma-1}k$. Bring (28) into (27) and we have

$$\begin{aligned}
&b_M^\sigma\tau_1^{1-\sigma}\Phi_{1u}+\beta\left(\frac{1}{2}b_H^\sigma+\frac{1}{2}b_L^\sigma-b_M^\sigma\right)\left[(1-\gamma)\,\tau_1^{1-\sigma}+\gamma\lambda_2\tau_2^{1-\sigma}+\gamma\,(1-\lambda_2)\,\tau_0^{1-\sigma}\right]\Phi_{1u}\\
&+\frac{\beta\gamma\lambda_2}{1-\beta}\left(b_M^\sigma\tau_2^{1-\sigma}\Phi_{1u}+\beta\left(\frac{1}{2}b_H^\sigma+\frac{1}{2}b_L^\sigma-b_M^\sigma\right)\tau_2^{1-\sigma}\Phi_{1u}\right)\\
&=(1-\beta+\beta\gamma\lambda_2)\left(b_M^\sigma\tau_1^{1-\sigma}\Phi_1+\frac{1}{2}\frac{\beta}{1-\beta}\left(b_H^\sigma+b_L^\sigma\right)\tau_1^{1-\sigma}\Phi_1\right)
\end{aligned} \tag{29}$$

Recall that $b_H=\varepsilon b_M$ and $b_L=(2-\varepsilon)\,b_M$. The above equation can be simplified as

$$\begin{aligned}
&\frac{\Phi_{1u}}{\Phi_1}\\
&=\frac{1+\frac{1}{2}\beta\left[b(\varepsilon)-2\right]+\beta\gamma\lambda_2\left(1+\frac{1}{2}\frac{\beta}{1-\beta}b(\varepsilon)\right)}{1+\frac{1}{2}\beta\left[b(\varepsilon)-2\right]\left[1-\gamma+\gamma\lambda_2\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}+\gamma\,(1-\lambda_2)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\right]+\beta\gamma\lambda_2\left(1+\frac{1}{2}\frac{\beta}{1-\beta}b(\varepsilon)\right)\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}}
\end{aligned} \tag{30}$$

Where $b(\varepsilon)=\varepsilon^\sigma+(2-\varepsilon)^\sigma$ and $b(\varepsilon)-2\geq 0$.

E Monotonicity of $\frac{\varphi_{1u}}{\varphi_1}$ regarding to λ_2

Recall that the monotonicity of $\frac{\varphi_{1u}}{\varphi_1}$ regarding to λ_2 is the same as that of $\frac{\varphi_{1u}^{\sigma-1}}{\varphi_1^{\sigma-1}}$ as σ is assumed to be greater than 1. Take the first derivative of $\frac{\varphi_{1u}^{\sigma-1}}{\varphi_1^{\sigma-1}}$ regarding to λ_2 and take the positive denominator to the left hand side.

$$
\begin{aligned}
&\frac{\partial \frac{\varphi_{1u}^{\sigma-1}}{\varphi_1^{\sigma-1}}}{\partial \lambda_2} (\cdot)^2 \\
=&\left(1+\frac{1}{2}\beta\left[b(\varepsilon)-2\right]\left[1-\gamma+\gamma\lambda_2\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}+\gamma\left(1-\lambda_2\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\right]+\beta\gamma\lambda_2\left(1+\frac{1}{2}\frac{\beta}{1-\beta}b(\varepsilon)\right)\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}\right) \\
&\cdot\beta\gamma\left(1+\frac{1}{2}\frac{\beta}{1-\beta}b(\varepsilon)\right) \\
&-\left(\frac{1}{2}\beta\left[b(\varepsilon)-2\right]\left[\gamma\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}-\gamma\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\right]+\beta\gamma\left(1+\frac{1}{2}\frac{\beta}{1-\beta}b(\varepsilon)\right)\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}\right) \\
&\cdot\left(1+\frac{1}{2}\beta\left[b(\varepsilon)-2\right]+\beta\gamma\lambda_2\left(1+\frac{1}{2}\frac{\beta}{1-\beta}b(\varepsilon)\right)\right) \\
=&\beta\gamma\left(1+\frac{1}{2}\frac{\beta}{1-\beta}b(\varepsilon)\right)\left[1-\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}\right] \\
&+\frac{1}{2}\beta\left[b(\varepsilon)-2\right]\beta\gamma\left(1+\frac{1}{2}\frac{\beta}{1-\beta}b(\varepsilon)\right)\left[1-\gamma+\gamma\lambda_2\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}+\gamma\left(1-\lambda_2\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}-\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}\right] \\
&+\frac{1}{2}\beta\gamma\left[b(\varepsilon)-2\right]\left(1+\frac{1}{2}\beta\left[b(\varepsilon)-2\right]+\beta\gamma\lambda_2\left(1+\frac{1}{2}\frac{\beta}{1-\beta}b(\varepsilon)\right)\right)\left[\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}-\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}\right] \\
\geq& 0
\end{aligned}
\tag{31}
$$

Where $(\cdot)^2$ is the denominator of the first derivative. The equality is taken as $\tau_2 = \tau_0 = \tau_1$. Therefore, $\frac{\varphi_{1u}}{\varphi_1}$ is an increasing function of λ_2.

F Proposition 2

Here, I would like to prove $\frac{\partial \frac{\varphi_{1u}}{\varphi_1}}{\partial \gamma} \geq 0$ is equivalent to $\frac{\varphi_{1u}}{\varphi_1} \geq 1$ and vice versa. Recall that $\frac{\partial \frac{\varphi_{1u}}{\varphi_1}}{\partial \gamma} \geq 0$ is equivalent to $\frac{\partial \frac{\varphi_{1u}^{\sigma-1}}{\varphi_1^{\sigma-1}}}{\partial \gamma} \geq 0$ and $\frac{\varphi_{1u}}{\varphi_1} \geq 1$ is equivalent to $\frac{\varphi_{1u}^{\sigma-1}}{\varphi_1^{\sigma-1}} \geq 1$ as $\sigma > 1$. We have

$$
\begin{aligned}
&\frac{\varphi_{1u}^{\sigma-1}}{\varphi_1^{\sigma-1}} \geq 1 \Leftrightarrow \\
&\frac{1}{2}\beta\left[b(\varepsilon)-2\right]\left[-\gamma+\gamma\lambda_2\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}+\gamma\left(1-\lambda_2\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\right]+\beta\gamma\lambda_2\left(1+\frac{1}{2}\frac{\beta}{1-\beta}b(\varepsilon)\right)\left[\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}-1\right] \\
&\leq 0
\end{aligned}
\tag{32}
$$

Take the first derivative of $\frac{\varphi_{1u}^{\sigma-1}}{\varphi_1^{\sigma-1}}$ regarding to γ and take the positive denominator to the left hand side.

$$\frac{\partial \frac{\varphi_{1u}^{\sigma-1}}{\varphi_1^{\sigma-1}}}{\partial\gamma}(\cdot)^2$$

$$=\left(1+\frac{1}{2}\beta\left[b(\varepsilon)-2\right]\left[1-\gamma+\gamma\lambda_2\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}+\gamma\left(1-\lambda_2\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\right]+\beta\gamma\lambda_2\left(1+\frac{1}{2}\frac{\beta}{1-\beta}b(\varepsilon)\right)\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}\right)$$
$$\cdot\beta\lambda_2\left(1+\frac{1}{2}\frac{\beta}{1-\beta}b(\varepsilon)\right)$$
$$-\left(\frac{1}{2}\beta\left[b(\varepsilon)-2\right]\left[-1+\lambda_2\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}+\left(1-\lambda_2\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\right]+\beta\lambda_2\left(1+\frac{1}{2}\frac{\beta}{1-\beta}b(\varepsilon)\right)\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}\right)$$
$$\cdot\left(1+\frac{1}{2}\beta\left[b(\varepsilon)-2\right]+\beta\gamma\lambda_2\left(1+\frac{1}{2}\frac{\beta}{1-\beta}b(\varepsilon)\right)\right)$$
$$=\left(1+\frac{1}{2}\beta\left[b(\varepsilon)-2\right]\right)$$
$$\cdot-\frac{1}{2}\beta\left[b(\varepsilon)-2\right]\left[-\gamma+\gamma\lambda_2\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}+\gamma\left(1-\lambda_2\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\right]-\beta\gamma\lambda_2\left(1+\frac{1}{2}\frac{\beta}{1-\beta}b(\varepsilon)\right)\left[\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}-1\right] \tag{33}$$

Where $(\cdot)^2$ is the denominator of the first derivative which is positive. The two above equations show that $\frac{\partial\frac{\varphi_{1u}}{\varphi_1}}{\partial\gamma}\geq 0$ is equivalent to $\frac{\varphi_{1u}}{\varphi_1}\geq 1$. The proof of the other statement is trivial.

G Proposition 3

Take the first derivative of $\frac{\varphi_{1u}^{\sigma-1}}{\varphi_1^{\sigma-1}}$ regarding to ε and take the positive denominator to the left hand side.

$$\frac{\partial \frac{\varphi_{1u}^{\sigma-1}}{\varphi_1^{\sigma-1}}}{\partial\varepsilon}(\cdot)^2$$

$$=\left(1+\frac{1}{2}\beta\left[b(\varepsilon)-2\right]\left[1-\gamma+\gamma\lambda_2\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}+\gamma\left(1-\lambda_2\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\right]+\beta\gamma\lambda_2\left(1+\frac{1}{2}\frac{\beta}{1-\beta}b(\varepsilon)\right)\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}\right)$$
$$\cdot\left(\frac{1}{2}\beta b'(\varepsilon)+\frac{1}{2}\frac{\beta}{1-\beta}\beta\gamma\lambda_2 b'(\varepsilon)\right)$$
$$-\left(\frac{1}{2}\beta b'(\varepsilon)\left[1-\gamma+\gamma\lambda_2\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}+\gamma\left(1-\lambda_2\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\right]+\frac{1}{2}\frac{\beta}{1-\beta}\beta\gamma\lambda_2 b'(\varepsilon)\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}\right)$$
$$\cdot\left(1+\frac{1}{2}\beta\left[b(\varepsilon)-2\right]+\beta\gamma\lambda_2\left(1+\frac{1}{2}\frac{\beta}{1-\beta}b(\varepsilon)\right)\right)$$
$$=\frac{1}{2}\beta\gamma\left(\gamma\lambda_2\frac{\beta}{1-\beta}+1\right)b'(\varepsilon)\left[1-\lambda_2\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}-\left(1-\lambda_2\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\right] \tag{34}$$

Where $b'(\varepsilon)=\sigma\varepsilon^{\sigma-1}-\sigma(2-\varepsilon)^{\sigma-1}>0$. Therefore, $\frac{\partial\frac{\varphi_{1u}}{\varphi_1}}{\partial\varepsilon}>0$ as $\lambda_2\tau_2^{-\sigma}+(1-\lambda_2)\tau_0^{-\sigma}<\tau_1^{-\sigma}$ and vice versa.

H Calculate $\frac{\varphi_{1u}}{\varphi_1}$ using fixed cost learning model

In this section, I consider a model with fixed cost and learning but without sunk cost. Define $\Pi_e(\varphi, \tau)$ as the expected value from exporting conditional on entry (entry means first-time exporting) with entry condition being τ and φ. $\Pi_e(\varphi, \tau_1)$ is

$$\begin{aligned}
\Pi_e(\varphi,\tau_1) =& b_M^\sigma \tau_1^{-\sigma}\Phi - f \\
&+ \gamma\lambda_2 \frac{\beta}{1-\beta}\left(\frac{1}{2}\mathbb{1}_{\varphi,\tau_2,b_H} b_H^\sigma + \frac{1}{2}\mathbb{1}_{\varphi,\tau_2,b_L} b_L^\sigma\right)\tau_2^{-\sigma}\Phi \\
&+ \gamma\lambda_2 \frac{\beta}{1-\beta}\left(-\frac{1}{2}\mathbb{1}_{\varphi,\tau_2,b_H} - \frac{1}{2}\mathbb{1}_{\varphi,\tau_2,b_L}\right) f \\
&+ \gamma\left(1-\lambda_2\right) \frac{\beta}{1-\beta}\left(\frac{1}{2}\mathbb{1}_{\varphi,\tau_0,b_H} b_H^\sigma + \frac{1}{2}\mathbb{1}_{\varphi,\tau_0,b_L} b_L^\sigma\right)\tau_0^{-\sigma}\Phi \\
&+ \gamma\left(1-\lambda_2\right) \frac{\beta}{1-\beta}\left(-\frac{1}{2}\mathbb{1}_{\varphi,\tau_0,b_H} - \frac{1}{2}\mathbb{1}_{\varphi,\tau_0,b_L}\right) f \\
&+ \left(1-\gamma\right)\beta\left(\frac{1}{2}\mathbb{1}_{\varphi,\tau_1,b_H} b_H^\sigma + \frac{1}{2}\mathbb{1}_{\varphi,\tau_1,b_L} b_L^\sigma\right)\tau_1^{-\sigma}\Phi \\
&+ \left(1-\gamma\right)\beta\left(-\frac{1}{2}\mathbb{1}_{\varphi,\tau_1,b_H} - \frac{1}{2}\mathbb{1}_{\varphi,\tau_1,b_L}\right) f \\
&+ \left(1-\gamma\right)\beta\Re\left(\varphi,\tau_1\right) \qquad (35) \\
=& b_M^\sigma \tau_1^{-\sigma}\Phi - f \\
&+ \frac{\gamma\lambda_2}{1-\left(1-\gamma\right)\beta}\frac{\beta}{1-\beta}\left(\frac{1}{2}\mathbb{1}_{\varphi,\tau_2,b_H} b_H^\sigma + \frac{1}{2}\mathbb{1}_{\varphi,\tau_2,b_L} b_L^\sigma\right)\tau_2^{-\sigma}\Phi \\
&+ \frac{\gamma\lambda_2}{1-\left(1-\gamma\right)\beta}\frac{\beta}{1-\beta}\left(-\frac{1}{2}\mathbb{1}_{\varphi,\tau_2,b_H} - \frac{1}{2}\mathbb{1}_{\varphi,\tau_2,b_L}\right) f \\
&+ \frac{\gamma\left(1-\lambda_2\right)}{1-\left(1-\gamma\right)\beta}\frac{\beta}{1-\beta}\left(\frac{1}{2}\mathbb{1}_{\varphi,\tau_0,b_H} b_H^\sigma + \frac{1}{2}\mathbb{1}_{\varphi,\tau_0,b_L} b_L^\sigma\right)\tau_0^{-\sigma}\Phi \\
&+ \frac{\gamma\left(1-\lambda_2\right)}{1-\left(1-\gamma\right)\beta}\frac{\beta}{1-\beta}\left(-\frac{1}{2}\mathbb{1}_{\varphi,\tau_0,b_H} - \frac{1}{2}\mathbb{1}_{\varphi,\tau_0,b_L}\right) f \\
&+ \frac{\left(1-\gamma\right)\beta}{1-\left(1-\gamma\right)\beta}\left(\frac{1}{2}\mathbb{1}_{\varphi,\tau_1,b_H} b_H^\sigma + \frac{1}{2}\mathbb{1}_{\varphi,\tau_1,b_L} b_L^\sigma\right)\tau_1^{-\sigma}\Phi \\
&+ \frac{\left(1-\gamma\right)\beta}{1-\left(1-\gamma\right)\beta}\left(-\frac{1}{2}\mathbb{1}_{\varphi,\tau_1,b_H} - \frac{1}{2}\mathbb{1}_{\varphi,\tau_1,b_L}\right) f
\end{aligned}$$

Where $\Phi = \varphi^{\sigma-1}k$. $\mathbb{1}_{\varphi,\tau,b}$ is the indicator of per period profit conditional on productivity φ, current trade policy realization τ and current belief b. $\mathbb{1}_{\varphi,\tau,b} = 1$ if $\pi(\varphi,\tau,b) = b^\sigma\tau^{-\sigma}\varphi^{\sigma-1}k - f > 0$. $\mathbb{1}_{\varphi,\tau,b} = 0$ if $\pi(\varphi,\tau,b) < 0$. I introduce indicator function since using per period fixed cost, per period profit can be negative and firms won't export with a negative profit after learning.

Firm φ enters under τ_1 with unconditional belief b_M (line 1). With probability $\gamma\lambda_2$, trade policy realization will be τ_2 in period 2 and it will be τ_2 forever. Therefore, we have the term of line 2 and 3. With probability $\gamma\left(1-\lambda_2\right)$, trade policy realization will be τ_0 in period 2 and it will be τ_0 forever. Therefore, we have the term of line 4 and 5. With probability $1-\gamma$, trade policy realization will be τ_1 in period 2 and we have the term of line 6 and 7. $\Re\left(\varphi,\tau_1\right)$ is the recursive term which means if trade realization in period 2 is τ_1, then in period 3, it will repeat the same process as that in period

2. The recursive term $\Re(\varphi, \tau_1)$ is

$$
\begin{aligned}
\Re(\varphi, \tau_1) = & \gamma\lambda_2 \frac{\beta}{1-\beta} \left(\frac{1}{2} \mathbb{1}_{\varphi,\tau_2,b_H} b_H^\sigma + \frac{1}{2} \mathbb{1}_{\varphi,\tau_2,b_L} b_L^\sigma \right) \tau_2^{-\sigma} \Phi \\
& + \gamma\lambda_2 \frac{\beta}{1-\beta} \left(-\frac{1}{2} \mathbb{1}_{\varphi,\tau_2,b_H} - \frac{1}{2} \mathbb{1}_{\varphi,\tau_2,b_L} \right) f \\
& + \gamma (1-\lambda_2) \frac{\beta}{1-\beta} \left(\frac{1}{2} \mathbb{1}_{\varphi,\tau_0,b_H} b_H^\sigma + \frac{1}{2} \mathbb{1}_{\varphi,\tau_0,b_L} b_L^\sigma \right) \tau_0^{-\sigma} \Phi \\
& + \gamma (1-\lambda_2) \frac{\beta}{1-\beta} \left(-\frac{1}{2} \mathbb{1}_{\varphi,\tau_0,b_H} - \frac{1}{2} \mathbb{1}_{\varphi,\tau_0,b_L} \right) f \\
& + (1-\gamma) \beta \left(\frac{1}{2} \mathbb{1}_{\varphi,\tau_1,b_H} b_H^\sigma + \frac{1}{2} \mathbb{1}_{\varphi,\tau_1,b_L} b_L^\sigma \right) \tau_1^{-\sigma} \Phi \\
& + (1-\gamma) \beta \left(-\frac{1}{2} \mathbb{1}_{\varphi,\tau_1,b_H} - \frac{1}{2} \mathbb{1}_{\varphi,\tau_1,b_L} \right) f \\
& + (1-\gamma) \beta \Re(\varphi, \tau_1)
\end{aligned} \tag{36}
$$

$\Pi_e(\varphi, \tau_2)$ is

$$
\begin{aligned}
\Pi_e(\varphi, \tau_2) = & b_M^\sigma \tau_2^{-\sigma} \Phi - f + \frac{\beta}{1-\beta} \left(\frac{1}{2} \mathbb{1}_{\varphi,\tau_2,b_H} b_H^\sigma + \frac{1}{2} \mathbb{1}_{\varphi,\tau_2,b_L} b_L^\sigma \right) \tau_2^{-\sigma} \Phi \\
& + \frac{\beta}{1-\beta} \left(-\frac{1}{2} \mathbb{1}_{\varphi,\tau_2,b_H} - \frac{1}{2} \mathbb{1}_{\varphi,\tau_2,b_L} \right) f
\end{aligned} \tag{37}
$$

Firm φ enters under τ_2 with unconditional belief b_M. Trade policy realization will be τ_2 forever. Similarly, $\Pi_e(\varphi, \tau_0)$ is

$$
\begin{aligned}
\Pi_e(\varphi, \tau_0) = & b_M^\sigma \tau_0^{-\sigma} \Phi - f + \frac{\beta}{1-\beta} \left(\frac{1}{2} \mathbb{1}_{\varphi,\tau_0,b_H} b_H^\sigma + \frac{1}{2} \mathbb{1}_{\varphi,\tau_0,b_L} b_L^\sigma \right) \tau_0^{-\sigma} \Phi \\
& + \frac{\beta}{1-\beta} \left(-\frac{1}{2} \mathbb{1}_{\varphi,\tau_0,b_H} - \frac{1}{2} \mathbb{1}_{\varphi,\tau_0,b_L} \right) f
\end{aligned} \tag{38}
$$

Firm φ enters under τ_0 with unconditional belief b_M. Trade policy realization will be τ_0 forever. Then we can write $\Pi_e(\varphi, \tau_1)$, $\Pi_e(\varphi, \tau_2)$, $\Pi_e(\varphi, \tau_0)$ recursively.

$$
\begin{aligned}
\Pi_e(\varphi, \tau_1) = & b_M^\sigma \tau_1^{-\sigma} \Phi - f + (1-\gamma) \beta \left(\frac{1}{2} \mathbb{1}_{\varphi,\tau_1,b_H} b_H^\sigma + \frac{1}{2} \mathbb{1}_{\varphi,\tau_1,b_L} b_L^\sigma - b_M^\sigma \right) \tau_1^{-\sigma} \Phi \\
& + (1-\gamma) \beta \left(-\frac{1}{2} \mathbb{1}_{\varphi,\tau_1,b_H} - \frac{1}{2} \mathbb{1}_{\varphi,\tau_1,b_L} + 1 \right) f \\
& + \gamma\lambda_2 \beta \left(\frac{1}{2} \mathbb{1}_{\varphi,\tau_2,b_H} b_H^\sigma + \frac{1}{2} \mathbb{1}_{\varphi,\tau_2,b_L} b_L^\sigma - b_M^\sigma \right) \tau_2^{-\sigma} \Phi \\
& + \gamma\lambda_2 \beta \left(-\frac{1}{2} \mathbb{1}_{\varphi,\tau_2,b_H} - \frac{1}{2} \mathbb{1}_{\varphi,\tau_2,b_L} + 1 \right) f \\
& + \gamma (1-\lambda_2) \beta \left(\frac{1}{2} \mathbb{1}_{\varphi,\tau_0,b_H} b_H^\sigma + \frac{1}{2} \mathbb{1}_{\varphi,\tau_0,b_L} b_L^\sigma - b_M^\sigma \right) \tau_0^{-\sigma} \Phi \\
& + \gamma (1-\lambda_2) \beta \left(-\frac{1}{2} \mathbb{1}_{\varphi,\tau_0,b_H} - \frac{1}{2} \mathbb{1}_{\varphi,\tau_0,b_L} + 1 \right) f \\
& + \beta E_{\tau_1} \Pi_e(\varphi, \tau') \\
= & f(\varphi, \tau_1) + \beta E_{\tau_1} \Pi_e(\varphi, \tau')
\end{aligned} \tag{39}
$$

$$
\begin{aligned}
\Pi_e(\varphi,\tau_2) =& b_M^\sigma \tau_2^{-\sigma}\Phi - f + \beta\left(\frac{1}{2}\mathbb{1}_{\varphi,\tau_2,b_H} b_H^\sigma + \frac{1}{2}\mathbb{1}_{\varphi,\tau_2,b_L} b_L^\sigma - b_M^\sigma\right)\tau_2^{-\sigma}\Phi \\
&+ \beta\left(-\frac{1}{2}\mathbb{1}_{\varphi,\tau_2,b_H} - \frac{1}{2}\mathbb{1}_{\varphi,\tau_2,b_L} + 1\right) f + \beta E_{\tau_2}\Pi_e(\varphi,\tau') \\
=& f(\varphi,\tau_2) + \beta E_{\tau_2}\Pi_e(\varphi,\tau') \\
=& f(\varphi,\tau_2) + \beta\Pi_e(\varphi,\tau_2)
\end{aligned} \tag{40}
$$

$$
\begin{aligned}
\Pi_e(\varphi,\tau_0) =& b_M^\sigma \tau_0^{-\sigma}\Phi - f + \beta\left(\frac{1}{2}\mathbb{1}_{\varphi,\tau_0,b_H} b_H^\sigma + \frac{1}{2}\mathbb{1}_{\varphi,\tau_0,b_L} b_L^\sigma - b_M^\sigma\right)\tau_0^{-\sigma}\Phi \\
&+ \beta\left(-\frac{1}{2}\mathbb{1}_{\varphi,\tau_0,b_H} - \frac{1}{2}\mathbb{1}_{\varphi,\tau_0,b_L} + 1\right) f + \beta E_{\tau_0}\Pi_e(\varphi,\tau') \\
=& f(\varphi,\tau_0) + \beta E_{\tau_0}\Pi_e(\varphi,\tau') \\
=& f(\varphi,\tau_0) + \beta\Pi_e(\varphi,\tau_0)
\end{aligned} \tag{41}
$$

$\Pi_e(\varphi,\tau_1)$, $\Pi_e(\varphi,\tau_2)$, $\Pi_e(\varphi,\tau_0)$ can be summarized as a following equation

$$\Pi_e(\varphi,\tau) = f(\varphi,\tau) + \beta E_\tau \Pi_e(\varphi,\tau') \tag{42}$$

Define $\Pi(\varphi,\tau)$ as the expected value under trade policy τ for a potential exporter φ. We have

$$\Pi(\varphi,\tau) = \max\left\{\Pi_e(\varphi,\tau), \beta E_\tau \Pi(\varphi,\tau')\right\} \tag{43}$$

$\Pi(\varphi,\tau)$ is the maximal value between entering in the current period $\Pi_e(\varphi,\tau)$ and waiting in the current period $\beta E_\tau \Pi(\varphi,\tau')$. Minus each side by $\Pi_e(\varphi,\tau)$ in (43) and bring (42) into (43), we have a following equation

$$\Pi(\varphi,\tau) - \Pi_e(\varphi,\tau) = \max\left\{0, \beta E_\tau\left[\Pi(\varphi,\tau') - \Pi_e(\varphi,\tau')\right] - f(\varphi,\tau)\right\} \tag{44}$$

$\Pi(\varphi,\tau) - \Pi_e(\varphi,\tau)$ is the value net of the profits of entering in the current period. If it's positive, firm φ will choose to wait in the current condition τ. If it's 0, firm φ will enter in the current period. Define $V(\varphi,\tau) = \Pi(\varphi,\tau) - \Pi_e(\varphi,\tau)$ which is the net value of waiting conditional on τ and φ and we have

$$V(\varphi,\tau) = \max\left\{0, \beta E_\tau V(\varphi,\tau') - f(\varphi,\tau)\right\} \tag{45}$$

I will focus on the entry threshold under τ_1 - φ_{1u}. For φ_{1u} firms, following condition should be satisfied.

$$\beta E_{\tau_1} V(\varphi_{1u},\tau') - f(\varphi_{1u},\tau_1) = 0 \tag{46}$$

The above condition implies that $V(\varphi_{1u},\tau_1) = 0$. Besides, I assume that $V(\varphi_{1u},\tau_0) = 0$ and $V(\varphi_{1u},\tau_2) > 0$ are satisfied, which means φ_{1u} firms prefer to enter under τ_0 and wait under τ_2.

Ideally, we want $V(\varphi,\tau)$ to be a decreasing function of φ and increasing function of τ.

(A) For any given φ, $V(\varphi,\tau)$ is an increasing function of τ?

(i) The problem is whether $f(\varphi,\tau)$ is a decreasing function of τ. We can easily observe that $f(\varphi,\tau_0) > f(\varphi,\tau_1)$. However, the relation between $f(\varphi,\tau_1)$ and $f(\varphi,\tau_2)$ is not obvious. $f(\varphi,\tau_1) > f(\varphi,\tau_2)$ for any φ if and only if ${\tau_1}^{-\sigma} - (1-\gamma)\beta{\tau_1}^{-\sigma} - \gamma\lambda_2\beta{\tau_2}^{-\sigma} - \gamma(1-\lambda_2)\beta{\tau_0}^{-\sigma} > (1-\beta){\tau_2}^{-\sigma}$.

(ii) Because of first-order dominance, $E_{\tau_0}V(\varphi,\tau') < E_{\tau_1}V(\varphi,\tau') < E_{\tau_2}V(\varphi,\tau')$ if $V(\varphi,\tau)$ increases in τ.

So property (i) such that $f(\varphi,\tau)$ is a decreasing function of τ is not always satisfied. Assume that property (i) is satisfied. By property (i) and (ii), if we start with a $V(\varphi,\tau)$ increasing in τ, the fixed point to this iteration also increases in τ.

(B) For any given τ, $V(\varphi,\tau)$ is a decreasing function of φ?

Using fixed cost, $f(\varphi,\tau_1)$ may not be an increasing function of φ. Therefore, we need to assume it. $f(\varphi,\tau_1)$ increases in φ if and only if ${\tau_1}^{-\sigma} - (1-\gamma)\beta{\tau_1}^{-\sigma} - \gamma\lambda_2\beta{\tau_2}^{-\sigma} - \gamma(1-\lambda_2)\beta{\tau_0}^{-\sigma} > 0$. We can easily observe that if $f(\varphi,\tau_1) > f(\varphi,\tau_2)$ for any φ is satisfied, then $f(\varphi,\tau_1)$ increasing in φ should also be satisfied.

If both (A) and (B) are satisfied, potential entrants will be more willing to choose to wait in the current period if their productivity φ is low and the current trade policy realization τ is high. In fact, we may be able to impose some weaker assumption. For example, $f(\varphi,\tau_1) > f(\varphi,\tau_2)$ for any φ is a too strong assumption. We don't need to consider the potential entrants that don't export in all three tariff states since their value is always 0. Therefore, a less restrictive condition is ${\tau_1}^{-\sigma} - (1-\gamma)\beta{\tau_1}^{-\sigma} - \gamma\lambda_2\beta{\tau_2}^{-\sigma} - \gamma(1-\lambda_2)\beta{\tau_0}^{-\sigma} + \frac{1}{2}\gamma(1-\lambda_2)\beta\frac{b_H^\sigma{\tau_0}^{-\sigma}\varphi^{\sigma-1}k-f}{b_M^\sigma\varphi^{\sigma-1}} > (1-\beta){\tau_2}^{-\sigma}$ with $\varphi^{\sigma-1} \geq \frac{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)f}{\left(b_M^\sigma+\frac{1}{2}\frac{\beta}{1-\beta}b_H^\sigma\right)k}\tau_0^\sigma$.

Following the above assumption, we can calculate $E_{\tau_1}V(\varphi_{1u},\tau')$.

$$\begin{aligned} E_{\tau_1}V(\varphi_{1u},\tau') =& \gamma\lambda_2 V(\varphi_{1u},\tau_2) \\ =& \gamma\lambda_2\left(\beta E_{\tau_2}V(\varphi_{1u},\tau') - f(\varphi_{1u},\tau_2)\right) \end{aligned} \tag{47}$$

We have

$$\begin{aligned} E_{\tau_2}V(\varphi_{1u},\tau') =& V(\varphi_{1u},\tau_2) \\ =& \beta E_{\tau_2}V(\varphi_{1u},\tau') - f(\varphi_{1u},\tau_2) \end{aligned} \tag{48}$$

From (48), we can solve $E_{\tau_2}V(\varphi_{1u},\tau')$

$$E_{\tau_2}V(\varphi_{1u},\tau') = -\frac{f(\varphi_{1u},\tau_2)}{1-\beta} \tag{49}$$

Bring (49) into (47) and we have

$$E_{\tau_1}V(\varphi_{1u},\tau') = -\frac{\gamma\lambda_2}{1-\beta}f(\varphi_{1u},\tau_2) \tag{50}$$

Bring (50) into (46) and we have

$$f_{\tau_1}(\varphi_{1u}) + \frac{\beta\gamma\lambda_2}{1-\beta}f(\varphi_{1u},\tau_2) = 0 \tag{51}$$

Recall that φ_1 is the entry threshold in a no TPU case and we have

$$b_M^\sigma \tau_1^{-\sigma} \Phi_1 - f + \frac{1}{2} \frac{\beta}{1-\beta} \left(b_H^\sigma \tau_1^{-\sigma} \Phi_1 - f \right) = 0 \tag{52}$$

Rearrange the above function and we have

$$f = \frac{\left(b_M^\sigma + \frac{1}{2} \frac{\beta}{1-\beta} b_H^\sigma \right) \tau_1^{-\sigma} \Phi_1}{1 + \frac{1}{2} \frac{\beta}{1-\beta}} \tag{53}$$

Bring (53) into (51) and we have

$$\begin{aligned}
& b_M^\sigma \tau_1^{-\sigma} \Phi_{1u} - \frac{\left(b_M^\sigma + \frac{1}{2} \frac{\beta}{1-\beta} b_H^\sigma \right) \tau_1^{-\sigma} \Phi_1}{1 + \frac{1}{2} \frac{\beta}{1-\beta}} \\
& + (1-\gamma) \beta \left(\frac{1}{2} \mathbb{1}_{\varphi_{1u}, \tau_1, b_H} b_H^\sigma + \frac{1}{2} \mathbb{1}_{\varphi_{1u}, \tau_1, b_L} b_L^\sigma - b_M^\sigma \right) \tau_1^{-\sigma} \Phi_{1u} \\
& + (1-\gamma) \beta \left(-\frac{1}{2} \mathbb{1}_{\varphi_{1u}, \tau_1, b_H} - \frac{1}{2} \mathbb{1}_{\varphi_{1u}, \tau_1, b_L} + 1 \right) \frac{\left(b_M^\sigma + \frac{1}{2} \frac{\beta}{1-\beta} b_H^\sigma \right) \tau_1^{-\sigma} \Phi_1}{1 + \frac{1}{2} \frac{\beta}{1-\beta}} \\
& + \gamma \lambda_2 \beta \left(\frac{1}{2} \mathbb{1}_{\varphi_{1u}, \tau_2, b_H} b_H^\sigma + \frac{1}{2} \mathbb{1}_{\varphi_{1u}, \tau_2, b_L} b_L^\sigma - b_M^\sigma \right) \tau_2^{-\sigma} \Phi_{1u} \\
& + \gamma \lambda_2 \beta \left(-\frac{1}{2} \mathbb{1}_{\varphi_{1u}, \tau_2, b_H} - \frac{1}{2} \mathbb{1}_{\varphi_{1u}, \tau_2, b_L} + 1 \right) \frac{\left(b_M^\sigma + \frac{1}{2} \frac{\beta}{1-\beta} b_H^\sigma \right) \tau_1^{-\sigma} \Phi_1}{1 + \frac{1}{2} \frac{\beta}{1-\beta}} \\
& + \gamma (1-\lambda_2) \beta \left(\frac{1}{2} \mathbb{1}_{\varphi_{1u}, \tau_0, b_H} b_H^\sigma + \frac{1}{2} \mathbb{1}_{\varphi_{1u}, \tau_0, b_L} b_L^\sigma - b_M^\sigma \right) \tau_0^{-\sigma} \Phi_{1u} \\
& + \gamma (1-\lambda_2) \beta \left(-\frac{1}{2} \mathbb{1}_{\varphi_{1u}, \tau_0, b_H} - \frac{1}{2} \mathbb{1}_{\varphi_{1u}, \tau_0, b_L} + 1 \right) \frac{\left(b_M^\sigma + \frac{1}{2} \frac{\beta}{1-\beta} b_H^\sigma \right) \tau_1^{-\sigma} \Phi_1}{1 + \frac{1}{2} \frac{\beta}{1-\beta}} \\
& + \frac{\beta \gamma \lambda_2}{1-\beta} b_M^\sigma \tau_2^{-\sigma} \Phi_{1u} - \frac{\beta \gamma \lambda_2}{1-\beta} \frac{\left(b_M^\sigma + \frac{1}{2} \frac{\beta}{1-\beta} b_H^\sigma \right) \tau_1^{-\sigma} \Phi_1}{1 + \frac{1}{2} \frac{\beta}{1-\beta}} \\
& + \frac{\beta \gamma \lambda_2}{1-\beta} \beta \left(\frac{1}{2} \mathbb{1}_{\varphi_{1u}, \tau_2, b_H} b_H^\sigma + \frac{1}{2} \mathbb{1}_{\varphi_{1u}, \tau_2, b_L} b_L^\sigma - b_M^\sigma \right) \tau_2^{-\sigma} \Phi_{1u} \\
& + \frac{\beta \gamma \lambda_2}{1-\beta} \beta \left(-\frac{1}{2} \mathbb{1}_{\varphi_{1u}, \tau_2, b_H} - \frac{1}{2} \mathbb{1}_{\varphi_{1u}, \tau_2, b_L} + 1 \right) \frac{\left(b_M^\sigma + \frac{1}{2} \frac{\beta}{1-\beta} b_H^\sigma \right) \tau_1^{-\sigma} \Phi_1}{1 + \frac{1}{2} \frac{\beta}{1-\beta}} \\
=& 0
\end{aligned} \tag{54}$$

$\pi(\varphi_{1u}, \tau_1, b_M) = b_M^\sigma \tau_1^{-\sigma} \Phi_{1u} - f < 0$ should be satisfied. If $\pi(\varphi_{1u}, \tau_1, b_M) > 0$, φ_{1u} firm will enter under τ_1 with a positive expected profit in the entry period and she will strictly prefer to enter under τ_1 which is a contradiction.[4] Therefore, we have $\mathbb{1}_{\varphi_{1u}, \tau_2, b_L} = 0$ and $\mathbb{1}_{\varphi_{1u}, \tau_1, b_L} = 0$. Moreover, since

[4] Recall that $\frac{1}{2} b_H^\sigma + \frac{1}{2} b_L^\sigma \geq b_M^\sigma$. Therefore, we have $\frac{1}{2} \mathbb{1}_{\varphi, \tau, b_H} \pi(\varphi, \tau, b_H) + \frac{1}{2} \mathbb{1}_{\varphi, \tau, b_L} \pi(\varphi, \tau, b_L) \geq \pi(\varphi, \tau, b_M)$ for any given φ and τ. The unconditional expectation of post-learning per period profit is always greater than per period profit with prior belief and demand learning is advantageous. If $\pi(\varphi_{1u}, \tau_1, b_M) > 0$, φ_{1u} firms can benefit from learning while won't suffer from a negative entry-period expected profit. In this case, φ_{1u} firms will strictly prefer to enter under τ_1, which is a contradiction.

φ_{1u} firm is willing to enter under τ_0, she should be able to make positive profit conditional on τ_0 and b_H and we have $\mathbb{1}_{\varphi_{1u},\tau_0,b_H} = 1$. For now, the value of $\mathbb{1}_{\varphi_{1u},\tau_2,b_H}$, $\mathbb{1}_{\varphi_{1u},\tau_1,b_H}$ and $\mathbb{1}_{\varphi_{1u},\tau_0,b_L}$ cannot be determined. Recall that $b_H = \varepsilon b_M$ and $b_L = (2-\varepsilon) b_M$. Simplify (54) and we have

$$\begin{aligned}
&\frac{\Phi_{1u}}{\Phi_1} - \frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^\sigma}{1+\frac{1}{2}\frac{\beta}{1-\beta}} \\
&+ (1-\gamma)\beta\left(\frac{1}{2}\mathbb{1}_{\varphi_{1u},\tau_1,b_H}\varepsilon^\sigma - 1\right)\frac{\Phi_{1u}}{\Phi_1} + (1-\gamma)\beta\left(-\frac{1}{2}\mathbb{1}_{\varphi_{1u},\tau_1,b_H} + 1\right)\frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^\sigma}{1+\frac{1}{2}\frac{\beta}{1-\beta}} \\
&+ \gamma\lambda_2\beta\left(\frac{1}{2}\mathbb{1}_{\varphi_{1u},\tau_2,b_H}\varepsilon^\sigma - 1\right)\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}\frac{\Phi_{1u}}{\Phi_1} + \gamma\lambda_2\beta\left(-\frac{1}{2}\mathbb{1}_{\varphi_{1u},\tau_2,b_H} + 1\right)\frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^\sigma}{1+\frac{1}{2}\frac{\beta}{1-\beta}} \\
&+ \gamma(1-\lambda_2)\beta\left(\frac{1}{2}\varepsilon^\sigma + \frac{1}{2}\mathbb{1}_{\varphi_{1u},\tau_0,b_L}(2-\varepsilon)^\sigma - 1\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\frac{\Phi_{1u}}{\Phi_1} \\
&+ \gamma(1-\lambda_2)\beta\left(\frac{1}{2} - \frac{1}{2}\mathbb{1}_{\varphi_{1u},\tau_0,b_L}\right)\frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^\sigma}{1+\frac{1}{2}\frac{\beta}{1-\beta}} \\
&+ \frac{\beta\gamma\lambda_2}{1-\beta}\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}\frac{\Phi_{1u}}{\Phi_1} - \frac{\beta\gamma\lambda_2}{1-\beta}\frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^\sigma}{1+\frac{1}{2}\frac{\beta}{1-\beta}} \\
&+ \frac{\beta\gamma\lambda_2}{1-\beta}\beta\left(\frac{1}{2}\mathbb{1}_{\varphi_{1u},\tau_2,b_H}\varepsilon^\sigma - 1\right)\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}\frac{\Phi_{1u}}{\Phi_1} + \frac{\beta\gamma\lambda_2}{1-\beta}\beta\left(-\frac{1}{2}\mathbb{1}_{\varphi_{1u},\tau_2,b_H} + 1\right)\frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^\sigma}{1+\frac{1}{2}\frac{\beta}{1-\beta}} \\
=&0
\end{aligned} \tag{55}$$

We can bring $\frac{\Phi_{1u}}{\Phi_1}$ on one side and bring the rest terms on the other side. Then we have

$$\frac{\Phi_{1u}}{\Phi_1} = \frac{numerator}{denominator} \tag{56}$$

Where *numerator* is

$$\begin{aligned}
numerator = &\left(1 - (1-\gamma)\beta\left(-\frac{1}{2}\mathbb{1}_{\varphi_{1u},\tau_1,b_H} + 1\right) - \gamma\lambda_2\beta\left(-\frac{1}{2}\mathbb{1}_{\varphi_{1u},\tau_2,b_H} + 1\right)\right. \\
&\left. -\gamma(1-\lambda_2)\beta\left(\frac{1}{2} - \frac{1}{2}\mathbb{1}_{\varphi_{1u},\tau_0,b_L}\right) + \frac{\beta\gamma\lambda_2}{1-\beta} - \frac{\beta\gamma\lambda_2}{1-\beta}\beta\left(-\frac{1}{2}\mathbb{1}_{\varphi_{1u},\tau_2,b_H} + 1\right)\right)\frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^\sigma}{1+\frac{1}{2}\frac{\beta}{1-\beta}}
\end{aligned} \tag{57}$$

And *denominator* is

$$\begin{aligned}
denominator = &1 + (1-\gamma)\beta\left(\frac{1}{2}\mathbb{1}_{\varphi_{1u},\tau_1,b_H}\varepsilon^\sigma - 1\right) + \gamma\lambda_2\beta\left(\frac{1}{2}\mathbb{1}_{\varphi_{1u},\tau_2,b_H}\varepsilon^\sigma - 1\right)\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma} \\
&+ \gamma(1-\lambda_2)\beta\left(\frac{1}{2}\varepsilon^\sigma + \frac{1}{2}\mathbb{1}_{\varphi_{1u},\tau_0,b_L}(2-\varepsilon)^\sigma - 1\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma} \\
&+ \frac{\beta\gamma\lambda_2}{1-\beta}\left(\frac{\tau_2}{\tau_1}\right)^{1-\sigma} + \frac{\beta\gamma\lambda_2}{1-\beta}\beta\left(\frac{1}{2}\mathbb{1}_{\varphi_{1u},\tau_2,b_H}\varepsilon^\sigma - 1\right)\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}
\end{aligned} \tag{58}$$

As we can see from the equations above, discussing if $\frac{\Phi_{1u}}{\Phi_1}$ is greater or less than 1 is not trivial. The formula of $\frac{\Phi_{1u}}{\Phi_1}$ depends on three indicators - $\mathbb{1}_{\varphi_{1u},\tau_1,b_H}$, $\mathbb{1}_{\varphi_{1u},\tau_2,b_H}$ and $\mathbb{1}_{\varphi_{1u},\tau_0,b_L}$. If we know the exact value of these three indicators, we can obtain a solution of $\frac{\Phi_{1u}}{\Phi_1}$ being expressed by exogenous parameters only. In order to obtain the exact solution of $\frac{\Phi_{1u}}{\Phi_1}$, we need to discuss the value of three indicators case by case. Moreover, we can start by discussing the value of $\mathbb{1}_{\varphi_{1u},\tau_1,b_H}$ firstly in order to simplify our discussion.

I 6 possible solutions of $\frac{\varphi_{1u}^{\sigma-1}}{\varphi_1^{\sigma-1}}$

For the 3 undetermined indicators, I will show that there are 6 different combinations in total.

I.1 $\mathbb{1}_{\varphi_{1u},\tau_1,b_H} = 0$

In this case, we assume that $\mathbb{1}_{\varphi_{1u},\tau_1,b_H} = 0$. φ_{1u} firms cannot make positive profit under τ_1 conditional on good belief b_H. As $\mathbb{1}_{\varphi_{1u},\tau_1,b_H} = 0$, $\mathbb{1}_{\varphi_{1u},\tau_2,b_H} = 0$ is also satisfied. Recall that $\mathbb{1}_{\varphi_1,\tau_1,b_H} = 1$ because φ_1 firms can make positive profit under τ_1 conditional on good belief b_H in the case without TPU. Therefore, if $\mathbb{1}_{\varphi_{1u},\tau_1,b_H} = 0$, $\varphi_{1u} < \varphi_1$ and there will be more entrants. $\mathbb{1}_{\varphi_{1u},\tau_1,b_H} = 0$ is equivalent to $b_H^\sigma \tau_1^{-\sigma}\varphi_{1u}^{\sigma-1}k - f \le 0$. From equation (53), we have $f = \frac{\left(b_M^\sigma + \frac{1}{2}\frac{\beta}{1-\beta}b_H^\sigma\right)\tau_1^{-\sigma}\Phi_1}{1+\frac{1}{2}\frac{\beta}{1-\beta}}$. Substitute f by $\frac{\left(b_M^\sigma + \frac{1}{2}\frac{\beta}{1-\beta}b_H^\sigma\right)\tau_1^{-\sigma}\Phi_1}{1+\frac{1}{2}\frac{\beta}{1-\beta}}$ and $b_H^\sigma \tau_1^{-\sigma}\varphi_{1u}^{\sigma-1}k - f \le 0$ is equivalent to

$$\frac{\Phi_{1u}}{\Phi_1} \le \frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^\sigma}{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)\varepsilon^\sigma} \tag{59}$$

Where $\Phi = \varphi^{\sigma-1}k$. We can easily observe that the right-hand side of the inequality $\frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^\sigma}{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)\varepsilon^\sigma}$ is less than 1. Therefore, $\mathbb{1}_{\varphi_{1u},\tau_1,b_H} = 0$ implies that $\frac{\Phi_{1u}}{\Phi_1} < 1$, which means there are more entrants in the uncertainty case. Assuming $\mathbb{1}_{\varphi_{1u},\tau_1,b_H} = 0$ implies that $\mathbb{1}_{\varphi_{1u},\tau_2,b_H} = 0$ and we only need to discuss the value of indicator $\mathbb{1}_{\varphi_{1u},\tau_0,b_L}$.

(1) $\mathbb{1}_{\varphi_{1u},\tau_0,b_L} = 0$

$\mathbb{1}_{\varphi_{1u},\tau_0,b_L} = 0$ is equivalent to $b_L^\sigma \tau_0^{-\sigma}\varphi_{1u}^{\sigma-1}k - f < 0$. Substitute f using equation (53) and $b_L^\sigma \tau_0^{-\sigma}\varphi_{1u}^{\sigma-1}k - f < 0$ is equivalent to

$$\frac{\Phi_{1u}}{\Phi_1} < \frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^\sigma}{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)(2-\varepsilon)^\sigma\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}} \tag{60}$$

In this case, φ_{1u} firms cannot make positive profit under τ_0 conditional on bad belief b_L. Since we have assumed that $\mathbb{1}_{\varphi_{1u},\tau_1,b_H} = 0$ and $\mathbb{1}_{\varphi_{1u},\tau_0,b_L} = 0$, the exact formula of $\frac{\Phi_{1u}}{\Phi_1}$ is

$$\frac{\Phi_{1u}}{\Phi_1} = \frac{\left[1-(1-\gamma)\beta - \frac{1}{2}\gamma(1-\lambda_2)\beta\right]\frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^\sigma}{1+\frac{1}{2}\frac{\beta}{1-\beta}}}{1-(1-\gamma)\beta + \gamma(1-\lambda_2)\beta\left(\frac{1}{2}\varepsilon^\sigma - 1\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}} \tag{61}$$

In order to guarantee that the above solution is feasible, two parametric constraints should be satisfied. The first constraint is obtained using (59) and (61) which is

$$\frac{\left[1-(1-\gamma)\beta - \frac{1}{2}\gamma(1-\lambda_2)\beta\right]\varepsilon^\sigma}{1-(1-\gamma)\beta - \frac{1}{2}\gamma(1-\lambda_2)\beta(2-\varepsilon^\sigma)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}} \le 1 \tag{62}$$

The second constraint is obtained using (60) and (61) which is

$$\frac{\left[1-(1-\gamma)\beta-\frac{1}{2}\gamma(1-\lambda_2)\beta\right](2-\varepsilon)^{\sigma}\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}}{1-(1-\gamma)\beta-\frac{1}{2}\gamma(1-\lambda_2)\beta(2-\varepsilon^{\sigma})\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}}<1 \tag{63}$$

(2) $\mathbb{1}_{\varphi_{1u},\tau_0,b_L}=1$

$\mathbb{1}_{\varphi_{1u},\tau_0,b_L}=1$ is equivalent to $b_L^{\sigma}\tau_0^{-\sigma}\varphi_{1u}^{\sigma-1}k-f>0$. Substitute f using equation (53) and $b_L^{\sigma}\tau_0^{-\sigma}\varphi_{1u}^{\sigma-1}k-f>0$ is equivalent to

$$\frac{\Phi_{1u}}{\Phi_1}>\frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^{\sigma}}{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)(2-\varepsilon)^{\sigma}\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}} \tag{64}$$

In this case, φ_{1u} firms can make positive profit under τ_0 conditional on bad belief b_L. Since we have assumed that $\mathbb{1}_{\varphi_{1u},\tau_1,b_H}=0$ and $\mathbb{1}_{\varphi_{1u},\tau_0,b_L}=1$, we can give the exact formula of $\frac{\Phi_{1u}}{\Phi_1}$

$$\frac{\Phi_{1u}}{\Phi_1}=\frac{[1-(1-\gamma)\beta]\frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^{\sigma}}{1+\frac{1}{2}\frac{\beta}{1-\beta}}}{1-(1-\gamma)\beta+\gamma(1-\lambda_2)\beta\left(\frac{1}{2}\varepsilon^{\sigma}+\frac{1}{2}(2-\varepsilon)^{\sigma}-1\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}} \tag{65}$$

In order to guarantee that the above solution is feasible, two parametric constraints should be satisfied. The first constraint is obtained using (59) and (65) which is

$$\frac{[1-(1-\gamma)\beta]\varepsilon^{\sigma}}{1-(1-\gamma)\beta+\gamma(1-\lambda_2)\beta\left(\frac{1}{2}\varepsilon^{\sigma}+\frac{1}{2}(2-\varepsilon)^{\sigma}-1\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}}\leq 1 \tag{66}$$

The second constraint is obtained using (64) and (65) which is

$$\frac{[1-(1-\gamma)\beta](2-\varepsilon)^{\sigma}\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}}{1-(1-\gamma)\beta+\gamma(1-\lambda_2)\beta\left(\frac{1}{2}\varepsilon^{\sigma}+\frac{1}{2}(2-\varepsilon)^{\sigma}-1\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}}>1 \tag{67}$$

I.2 $\mathbb{1}_{\varphi_{1u},\tau_1,b_H}=1$

In this case, we assume that $\mathbb{1}_{\varphi_{1u},\tau_1,b_H}=1$. φ_{1u} firms can make positive profit under τ_1 conditional on good belief b_H. We cannot tell directly the relation between φ_{1u} and φ_1 and we need to discuss 4 different cases. $\mathbb{1}_{\varphi_{1u},\tau_1,b_H}=1$ is equivalent to $b_H^{\sigma}\tau_1^{-\sigma}\varphi_{1u}^{\sigma-1}k-f>0$ which is also equivalent to

$$\frac{\Phi_{1u}}{\Phi_1}>\frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^{\sigma}}{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)\varepsilon^{\sigma}} \tag{68}$$

(1) $\mathbb{1}_{\varphi_{1u},\tau_2,b_H}=0$ **and** $\mathbb{1}_{\varphi_{1u},\tau_0,b_L}=0$

In the first sub case, we assume that $\mathbb{1}_{\varphi_{1u},\tau_2,b_H}=0$ and $\mathbb{1}_{\varphi_{1u},\tau_0,b_L}=0$. φ_{1u} firms cannot make positive profit under τ_2 conditional on good belief b_H. Besides, they cannot make positive profit under τ_0 conditional on bad belief b_L either. $\mathbb{1}_{\varphi_{1u},\tau_2,b_H}=0$ is equivalent to

$$\frac{\Phi_{1u}}{\Phi_1}<\frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^{\sigma}}{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)\varepsilon^{\sigma}\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}} \tag{69}$$

And $\mathbb{1}_{\varphi_{1u},\tau_0,b_L} = 0$ is equivalent to

$$\frac{\Phi_{1u}}{\Phi_1} < \frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^\sigma}{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)(2-\varepsilon)^\sigma\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}} \tag{70}$$

Since we have assumed that $\mathbb{1}_{\varphi_{1u},\tau_1,b_H} = 1$, $\mathbb{1}_{\varphi_{1u},\tau_2,b_H} = 0$ and $\mathbb{1}_{\varphi_{1u},\tau_0,b_L} = 0$, we can give the exact formula of $\frac{\Phi_{1u}}{\Phi_1}$

$$\frac{\Phi_{1u}}{\Phi_1} = \frac{\left[1-\frac{1}{2}(1-\gamma)\beta-\frac{1}{2}\gamma(1-\lambda_2)\beta\right]\frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^\sigma}{1+\frac{1}{2}\frac{\beta}{1-\beta}}}{1-\left[\frac{1}{2}(1-\gamma)\beta+\frac{1}{2}\gamma(1-\lambda_2)\beta\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\right](2-\varepsilon^\sigma)} \tag{71}$$

In order to guarantee that the above solution is feasible, three parametric constraints should be satisfied. The first constraint is obtained using (68) and (71) which is

$$\frac{\left[1-\frac{1}{2}(1-\gamma)\beta-\frac{1}{2}\gamma(1-\lambda_2)\beta\right]\varepsilon^\sigma}{1-\left[\frac{1}{2}(1-\gamma)\beta+\frac{1}{2}\gamma(1-\lambda_2)\beta\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\right](2-\varepsilon^\sigma)} > 1 \tag{72}$$

The second constraint is obtained using (69) and (71) which is

$$\frac{\left[1-\frac{1}{2}(1-\gamma)\beta-\frac{1}{2}\gamma(1-\lambda_2)\beta\right]\varepsilon^\sigma\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}}{1-\left[\frac{1}{2}(1-\gamma)\beta+\frac{1}{2}\gamma(1-\lambda_2)\beta\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\right](2-\varepsilon^\sigma)} < 1 \tag{73}$$

The third constraint is obtained using (70) and (71) which is

$$\frac{\left[1-\frac{1}{2}(1-\gamma)\beta-\frac{1}{2}\gamma(1-\lambda_2)\beta\right](2-\varepsilon)^\sigma\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}}{1-\left[\frac{1}{2}(1-\gamma)\beta+\frac{1}{2}\gamma(1-\lambda_2)\beta\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\right](2-\varepsilon^\sigma)} < 1 \tag{74}$$

(2) $\mathbb{1}_{\varphi_{1u},\tau_2,b_H} = 1$ and $\mathbb{1}_{\varphi_{1u},\tau_0,b_L} = 0$

In the second sub case, we assume that $\mathbb{1}_{\varphi_{1u},\tau_2,b_H} = 1$ and $\mathbb{1}_{\varphi_{1u},\tau_0,b_L} = 0$. φ_{1u} firms can make positive profit under τ_2 conditional on good belief b_H while they cannot make positive profit under τ_0 conditional on bad belief b_L. $\mathbb{1}_{\varphi_{1u},\tau_2,b_H} = 1$ is equivalent to

$$\frac{\Phi_{1u}}{\Phi_1} > \frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^\sigma}{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)\varepsilon^\sigma\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}} \tag{75}$$

$\mathbb{1}_{\varphi_{1u},\tau_0,b_L} = 0$ is equivalent to

$$\frac{\Phi_{1u}}{\Phi_1} < \frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^\sigma}{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)(2-\varepsilon)^\sigma\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}} \tag{76}$$

Since we have assumed that $\mathbb{1}_{\varphi_{1u},\tau_1,b_H}=1$, $\mathbb{1}_{\varphi_{1u},\tau_2,b_H}=1$ and $\mathbb{1}_{\varphi_{1u},\tau_0,b_L}=0$, we can give the exact formula of $\frac{\Phi_{1u}}{\Phi_1}$

$$\frac{\Phi_{1u}}{\Phi_1}=\frac{\left(1-\frac{1}{2}\beta\right)\left(1+\frac{\beta\gamma\lambda_2}{1-\beta}\right)\frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^\sigma}{1+\frac{1}{2}\frac{\beta}{1-\beta}}}{1+\left[(1-\gamma)\beta+\gamma(1-\lambda_2)\beta\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\right]\left(\frac{1}{2}\varepsilon^\sigma-1\right)+\frac{1}{2}\varepsilon^\sigma\frac{\beta\gamma\lambda_2}{1-\beta}\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}} \tag{77}$$

In order to guarantee that the above solution is feasible, three parametric constraints should be satisfied. The first constraint is obtained using (68) and (77) which is

$$\frac{\left(1-\frac{1}{2}\beta\right)\left(1+\frac{\beta\gamma\lambda_2}{1-\beta}\right)\varepsilon^\sigma}{1+\left[(1-\gamma)\beta+\gamma(1-\lambda_2)\beta\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\right]\left(\frac{1}{2}\varepsilon^\sigma-1\right)+\frac{1}{2}\varepsilon^\sigma\frac{\beta\gamma\lambda_2}{1-\beta}\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}}>1 \tag{78}$$

The second constraint is obtained using (75) and (77) which is

$$\frac{\left(1-\frac{1}{2}\beta\right)\left(1+\frac{\beta\gamma\lambda_2}{1-\beta}\right)\varepsilon^\sigma\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}}{1+\left[(1-\gamma)\beta+\gamma(1-\lambda_2)\beta\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\right]\left(\frac{1}{2}\varepsilon^\sigma-1\right)+\frac{1}{2}\varepsilon^\sigma\frac{\beta\gamma\lambda_2}{1-\beta}\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}}>1 \tag{79}$$

The third constraint is obtained using (76) and (77) which is

$$\frac{\left(1-\frac{1}{2}\beta\right)\left(1+\frac{\beta\gamma\lambda_2}{1-\beta}\right)(2-\varepsilon)^\sigma\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}}{1+\left[(1-\gamma)\beta+\gamma(1-\lambda_2)\beta\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\right]\left(\frac{1}{2}\varepsilon^\sigma-1\right)+\frac{1}{2}\varepsilon^\sigma\frac{\beta\gamma\lambda_2}{1-\beta}\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}}<1 \tag{80}$$

(3) $\mathbb{1}_{\varphi_{1u},\tau_2,b_H}=0$ and $\mathbb{1}_{\varphi_{1u},\tau_0,b_L}=1$

In the third sub case, we assume that $\mathbb{1}_{\varphi_{1u},\tau_2,b_H}=0$ and $\mathbb{1}_{\varphi_{1u},\tau_0,b_L}=1$. φ_{1u} firms cannot make positive profit under τ_2 conditional on good belief b_H while they can make positive profit under τ_0 conditional on bad belief b_L. $\mathbb{1}_{\varphi_{1u},\tau_2,b_H}=0$ is equivalent to

$$\frac{\Phi_{1u}}{\Phi_1}<\frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^\sigma}{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)\varepsilon^\sigma\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}} \tag{81}$$

$\mathbb{1}_{\varphi_{1u},\tau_0,b_L}=1$ is equivalent to

$$\frac{\Phi_{1u}}{\Phi_1}>\frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^\sigma}{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)(2-\varepsilon)^\sigma\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}} \tag{82}$$

Since we have assumed that $\mathbb{1}_{\varphi_{1u},\tau_1,b_H}=1$, $\mathbb{1}_{\varphi_{1u},\tau_2,b_H}=0$ and $\mathbb{1}_{\varphi_{1u},\tau_0,b_L}=1$, we can give the exact formula of $\frac{\Phi_{1u}}{\Phi_1}$

$$\frac{\Phi_{1u}}{\Phi_1}=\frac{\left[1-\frac{1}{2}(1-\gamma)\beta\right]\frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^\sigma}{1+\frac{1}{2}\frac{\beta}{1-\beta}}}{1-\frac{1}{2}(1-\gamma)\beta\left(2-\varepsilon^\sigma\right)+\gamma(1-\lambda_2)\beta\left(\frac{1}{2}\varepsilon^\sigma+\frac{1}{2}(2-\varepsilon)^\sigma-1\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}} \tag{83}$$

In order to guarantee that the above solution is feasible, three parametric constraints should be satisfied. The first constraint is obtained using (68) and (83) which is

$$\frac{\left[1-\frac{1}{2}(1-\gamma)\beta\right]\varepsilon^{\sigma}}{1-\frac{1}{2}(1-\gamma)\beta\left(2-\varepsilon^{\sigma}\right)+\gamma(1-\lambda_2)\beta\left(\frac{1}{2}\varepsilon^{\sigma}+\frac{1}{2}(2-\varepsilon)^{\sigma}-1\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}}>1 \tag{84}$$

The second constraint is obtained using (81) and (83) which is

$$\frac{\left[1-\frac{1}{2}(1-\gamma)\beta\right]\varepsilon^{\sigma}\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}}{1-\frac{1}{2}(1-\gamma)\beta\left(2-\varepsilon^{\sigma}\right)+\gamma(1-\lambda_2)\beta\left(\frac{1}{2}\varepsilon^{\sigma}+\frac{1}{2}(2-\varepsilon)^{\sigma}-1\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}}<1 \tag{85}$$

The third constraint is obtained using (82) and (83) which is

$$\frac{\left[1-\frac{1}{2}(1-\gamma)\beta\right](2-\varepsilon)^{\sigma}\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}}{1-\frac{1}{2}(1-\gamma)\beta\left(2-\varepsilon^{\sigma}\right)+\gamma(1-\lambda_2)\beta\left(\frac{1}{2}\varepsilon^{\sigma}+\frac{1}{2}(2-\varepsilon)^{\sigma}-1\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}}>1 \tag{86}$$

(4) $\mathbb{1}_{\varphi_{1u},\tau_2,b_H}=1$ **and** $\mathbb{1}_{\varphi_{1u},\tau_0,b_L}=1$

In the fourth sub case, we assume that $\mathbb{1}_{\varphi_{1u},\tau_2,b_H}=1$ and $\mathbb{1}_{\varphi_{1u},\tau_0,b_L}=1$. φ_{1u} firms can make positive profit under τ_2 conditional on good belief b_H. Besides, they can also make positive profit under τ_0 conditional on bad belief b_L. $\mathbb{1}_{\varphi_{1u},\tau_2,b_H}=1$ is equivalent to

$$\frac{\Phi_{1u}}{\Phi_1}>\frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^{\sigma}}{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)\varepsilon^{\sigma}\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}} \tag{87}$$

$\mathbb{1}_{\varphi_{1u},\tau_0,b_L}=1$ is equivalent to

$$\frac{\Phi_{1u}}{\Phi_1}>\frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^{\sigma}}{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)(2-\varepsilon)^{\sigma}\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}} \tag{88}$$

Since we have assumed that $\mathbb{1}_{\varphi_{1u},\tau_1,b_H}=1$, $\mathbb{1}_{\varphi_{1u},\tau_2,b_H}=1$ and $\mathbb{1}_{\varphi_{1u},\tau_0,b_L}=1$, we can give the exact formula of $\frac{\Phi_{1u}}{\Phi_1}$

$$\frac{\Phi_{1u}}{\Phi_1}=\frac{\left[1-\frac{1}{2}(1-\gamma)\beta+\frac{1}{2}\frac{\beta\gamma\lambda_2}{1-\beta}\right]\frac{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^{\sigma}}{1+\frac{1}{2}\frac{\beta}{1-\beta}}}{1-\frac{1}{2}(1-\gamma)\beta\left(2-\varepsilon^{\sigma}\right)+\frac{1}{2}\varepsilon^{\sigma}\frac{\beta\gamma\lambda_2}{1-\beta}\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}+\gamma(1-\lambda_2)\beta\left(\frac{1}{2}\varepsilon^{\sigma}+\frac{1}{2}(2-\varepsilon)^{\sigma}-1\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}} \tag{89}$$

In order to guarantee that the above solution is feasible, three parametric constraints should be satisfied. The first constraint is obtained using (68) and (89) which is

$$\frac{\left[1-\frac{1}{2}(1-\gamma)\beta+\frac{1}{2}\frac{\beta\gamma\lambda_2}{1-\beta}\right]\varepsilon^{\sigma}}{1-\frac{1}{2}(1-\gamma)\beta\left(2-\varepsilon^{\sigma}\right)+\frac{1}{2}\varepsilon^{\sigma}\frac{\beta\gamma\lambda_2}{1-\beta}\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}+\gamma(1-\lambda_2)\beta\left(\frac{1}{2}\varepsilon^{\sigma}+\frac{1}{2}(2-\varepsilon)^{\sigma}-1\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}}>1 \tag{90}$$

The second constraint is obtained using (87) and (89) which is

$$\frac{\left[1-\frac{1}{2}(1-\gamma)\beta+\frac{1}{2}\frac{\beta\gamma\lambda_2}{1-\beta}\right]\varepsilon^{\sigma}\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}}{1-\frac{1}{2}(1-\gamma)\beta\left(2-\varepsilon^{\sigma}\right)+\frac{1}{2}\varepsilon^{\sigma}\frac{\beta\gamma\lambda_2}{1-\beta}\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}+\gamma(1-\lambda_2)\beta\left(\frac{1}{2}\varepsilon^{\sigma}+\frac{1}{2}(2-\varepsilon)^{\sigma}-1\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}}>1 \tag{91}$$

The third constraint is obtained using (88) and (89) which is

$$\frac{\left[1-\frac{1}{2}(1-\gamma)\beta+\frac{1}{2}\frac{\beta\gamma\lambda_2}{1-\beta}\right](2-\varepsilon)^{\sigma}\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}}{1-\frac{1}{2}(1-\gamma)\beta\left(2-\varepsilon^{\sigma}\right)+\frac{1}{2}\varepsilon^{\sigma}\frac{\beta\gamma\lambda_2}{1-\beta}\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}+\gamma(1-\lambda_2)\beta\left(\frac{1}{2}\varepsilon^{\sigma}+\frac{1}{2}(2-\varepsilon)^{\sigma}-1\right)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}}>1 \tag{92}$$

J Lemma 3

J.1 φ_1 firms are willing to enter under τ_1

In this case, φ_1 firms are assumed to be willing to enter under τ_1 as TPU is imposed. The value of entering under τ_1 - $\Pi_e(\varphi_1,\tau_1)$ is

$$\begin{aligned}\Pi_e(\varphi_1,\tau_1)=&\pi\left(\varphi_1,\tau_1,b_M\right)+\frac{1}{2}\frac{(1-\gamma)\beta}{1-(1-\gamma)\beta}\pi\left(\varphi_1,\tau_1,b_H\right)\\&+\frac{1}{2}\frac{\gamma\lambda_2}{1-(1-\gamma)\beta}\frac{\beta}{1-\beta}\tilde{\pi}\left(\varphi_1,\tau_2,b_H\right)\\&+\frac{1}{2}\frac{\gamma\left(1-\lambda_2\right)}{1-(1-\gamma)\beta}\frac{\beta}{1-\beta}\Big(\pi\left(\varphi_1,\tau_0,b_H\right)+\tilde{\pi}\left(\varphi_1,\tau_0,b_L\right)\Big)\end{aligned} \tag{93}$$

The value of waiting under τ_1 - $\beta E_{\tau_1}\Pi(\varphi_1,\tau')$ is

$$\begin{aligned}&\beta E_{\tau_1}\Pi(\varphi_1,\tau')\\=&(1-\gamma)\beta\Pi_e(\varphi_1,\tau_1)+\gamma(1-\lambda_2)\beta\Pi_e(\varphi_1,\tau_0)\\=&(1-\gamma)\beta\Pi_e(\varphi_1,\tau_1)+\gamma(1-\lambda_2)\beta\left[\pi\left(\varphi_1,\tau_0,b_M\right)+\frac{1}{2}\frac{\beta}{1-\beta}\Big(\pi\left(\varphi_1,\tau_0,b_H\right)+\tilde{\pi}\left(\varphi_1,\tau_0,b_L\right)\Big)\right]\end{aligned} \tag{94}$$

Recall that $\pi\left(\varphi_1,\tau_1,b_M\right)+\frac{1}{2}\frac{\beta}{1-\beta}\pi\left(\varphi_1,\tau_1,b_H\right)=0$. Therefore, the net difference between the value of entering and the value of waiting under τ_1 - $Diff\left(\varphi_1,\tau_1\right)=\Pi_e(\varphi_1,\tau_1)-\beta E_{\tau_1}\Pi(\varphi_1,\tau')$ is

$$\begin{aligned}&Diff\left(\varphi_1,\tau_1\right)\\=&\Big(1-(1-\gamma)\beta\Big)\Pi_e(\varphi_1,\tau_1)-\gamma(1-\lambda_2)\beta\Pi_e(\varphi_1,\tau_0)\\=&-\frac{1}{2}\gamma\frac{\beta}{1-\beta}\pi\left(\varphi_1,\tau_1,b_H\right)+\frac{1}{2}\gamma\lambda_2\frac{\beta}{1-\beta}\tilde{\pi}\left(\varphi_1,\tau_2,b_H\right)\\&+\frac{1}{2}\gamma(1-\lambda_2)\beta\Big(\pi\left(\varphi_1,\tau_0,b_H\right)+\tilde{\pi}\left(\varphi_1,\tau_0,b_L\right)-2\pi\left(\varphi_1,\tau_0,b_M\right)\Big)\\=&\frac{1}{2}\gamma\beta Sdiff\left(\varphi_1,\tau_1\right)\end{aligned} \tag{95}$$

J.2 φ_1 firms are willing to wait under τ_1

In this case, φ_1 firms are assumed to be willing to wait under τ_1 as TPU is imposed. The value of waiting under τ_1 - $\beta E_{\tau_1}\Pi(\varphi_1, \tau')$ is

$$\begin{aligned}
&\beta E_{\tau_1}\Pi(\varphi_1, \tau') \\
=&\frac{\gamma(1-\lambda_2)\beta}{1-\beta(1-\gamma)}\Pi_e(\varphi_1, \tau_0) \\
=&\frac{\gamma(1-\lambda_2)\beta}{1-\beta(1-\gamma)}\left[\pi\left(\varphi_1, \tau_0, b_M\right) + \frac{1}{2}\frac{\beta}{1-\beta}\Big(\pi\left(\varphi_1, \tau_0, b_H\right) + \tilde{\pi}\left(\varphi_1, \tau_0, b_L\right)\Big)\right]
\end{aligned} \tag{96}$$

Therefore, the net difference between the value of entering and the value of waiting under τ_1 - $Diff\left(\varphi_1, \tau_1\right) = \Pi_e(\varphi_1, \tau_1) - \beta E_{\tau_1}\Pi(\varphi_1, \tau')$ is

$$\begin{aligned}
&Diff\left(\varphi_1, \tau_1\right) \\
=&\Pi_e(\varphi_1, \tau_1) - \frac{\gamma(1-\lambda_2)\beta}{1-\beta(1-\gamma)}\Pi_e(\varphi_1, \tau_0) \\
=&\frac{1}{2}\frac{\gamma\beta}{1-\beta(1-\gamma)}Sdiff\left(\varphi_1, \tau_1\right)
\end{aligned} \tag{97}$$

K $Sdiff\left(\varphi_1, \tau_1\right)$ as a function of δ

Substitute τ_2 by $\delta\tau_1$, τ_0 by $\frac{1-\lambda_2\delta}{1-\lambda_2}\tau_1$, b_H by εb_M, b_L by $\left(2-\varepsilon\right)b_M$ and $\varphi_1^{\sigma-1}$ by $\frac{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)f}{\left(b_M^\sigma + \frac{1}{2}\frac{\beta}{1-\beta}b_H^\sigma\right)k}\tau_1^\sigma$. In addition, divide $Sdiff\left(\varphi_1, \tau_1\right)$ by f and we have

$$\begin{aligned}
&\frac{Sdiff\left(\varphi_1, \tau_1\right)}{f} = \\
&-\frac{B\varepsilon^\sigma - 1}{1-\beta} + \frac{\lambda_2}{1-\beta}\max\left\{B\varepsilon^\sigma\delta^{-\sigma} - 1, 0\right\} \\
&+\left(1-\lambda_2\right)\left(B\varepsilon^\sigma\left(\frac{1-\lambda_2\delta}{1-\lambda_2}\right)^{-\sigma} + \max\left\{B\left(2-\varepsilon\right)^\sigma\left(\frac{1-\lambda_2\delta}{1-\lambda_2}\right)^{-\sigma} - 1, 0\right\} - 2B\left(\frac{1-\lambda_2\delta}{1-\lambda_2}\right)^{-\sigma} + 1\right)
\end{aligned} \tag{98}$$

Where $B = \frac{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)}{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^\sigma}$. As $\delta \to 1$, $Sdiff\left(\varphi_1, \tau_1\right) \to 0$ and we go back to the case without TPU. There exist two zero profit cutoffs δ_{τ_2} and δ_{τ_0}. As $\delta < \delta_{\tau_2}$, $\mathbb{1}_{\varphi_1, \tau_2, b_H} = 1$ and $\max\left\{B\varepsilon^\sigma\delta^{-\sigma} - 1, 0\right\} = B\varepsilon^\sigma\delta^{-\sigma} - 1 > 0$. Otherwise $\mathbb{1}_{\varphi_1, \tau_2, b_H} = 0$. As $\delta > \delta_{\tau_0}$, $\mathbb{1}_{\varphi_1, \tau_0, b_L} = 1$ and $\max\left\{B\left(2-\varepsilon\right)^\sigma\left(\frac{1-\lambda_2\delta}{1-\lambda_2}\right)^{-\sigma} - 1, 0\right\} > 0$. Otherwise $\mathbb{1}_{\varphi_1, \tau_0, b_L} = 0$. We have $\delta_{\tau_2} = B^{\frac{1}{\sigma}}\varepsilon$ and $\delta_{\tau_0} = \frac{1-\lambda_2}{\lambda_2}\left[\frac{1}{1-\lambda_2} - B^{\frac{1}{\sigma}}\left(2-\varepsilon\right)\right]$. In addition, since τ_0 cannot go below 1, there exists a maximum of δ - δ_{max} which equals to $\frac{1-\lambda_2}{\lambda_2}\left(\frac{1}{1-\lambda_2} - \frac{1}{\tau_1}\right)$. Without loss of generality, I assume $f = 1$ in the current section.

K.1 $B(\varphi_1, \tau_2) = \frac{\lambda_2}{1-\beta}\max\left\{B\varepsilon^\sigma\delta^{-\sigma} - 1, 0\right\}$

$B(\varphi_1, \tau_2)$ is a function of δ and take the first derivative

$$\frac{\partial B(\varphi_1, \tau_2)}{\partial\delta} = -\mathbb{1}_{\varphi_1, \tau_2, b_H} * \sigma\frac{\lambda_2}{1-\beta}B\varepsilon^\sigma\delta^{-\sigma-1} \leq 0 \tag{99}$$

Take the second derivative

$$\frac{\partial^2 B(\varphi_1,\tau_2)}{\partial\delta^2} = \mathbb{1}_{\varphi_1,\tau_2,b_H} * \sigma(\sigma+1)\frac{\lambda_2}{1-\beta}B\varepsilon^\sigma\delta^{-\sigma-2} \geq 0 \tag{100}$$

As δ increases from 1 to $\delta^-_{\tau_2}$, $B(\varphi_1,\tau_2) = \frac{\lambda_2}{1-\beta}(B\varepsilon^\sigma\delta^{-\sigma}-1)$ which is a decreasing function of δ. In this case, φ_1 firms can still make positive profit under τ_2 conditional on good belief b_H. However, as δ increases, this profit becomes smaller and smaller. As δ increases from $\delta^-_{\tau_2}$ to $\delta^+_{\tau_2}$, $\mathbb{1}_{\varphi_1,\tau_2,b_H}$ turns from 1 to 0 and $B(\varphi_1,\tau_2) = 0$. In this case, exporting under τ_2 conditional on good belief b_H turns to nonprofitable and the profit is bounded at 0. Moreover, there will be an upward jump of the first derivative of $B(\varphi_1,\tau_2)$ (from negative to 0). As δ increases from $\delta^+_{\tau_2}$, $\mathbb{1}_{\varphi_1,\tau_2,b_H} = 0$ and $B(\varphi_1,\tau_2) = 0$. We can also see that the second derivative is always non-negative.

K.2 $G(\varphi_1,\tau_0) = (1-\lambda_2)\left(B\varepsilon^\sigma\left(\frac{1-\lambda_2\delta}{1-\lambda_2}\right)^{-\sigma} + \max\left\{B(2-\varepsilon)^\sigma\left(\frac{1-\lambda_2\delta}{1-\lambda_2}\right)^{-\sigma}-1,0\right\} - 2B\left(\frac{1-\lambda_2\delta}{1-\lambda_2}\right)^{-\sigma}+1\right)$

Take the first derivative

$$\frac{\partial G(\varphi_1,\tau_0)}{\partial\delta} = \sigma\lambda_2 B\left(\varepsilon^\sigma + \mathbb{1}_{\varphi_1,\tau_0,b_L} * (2-\varepsilon)^\sigma - 2\right)\left(\frac{1-\lambda_2\delta}{1-\lambda_2}\right)^{-\sigma-1} \tag{101}$$

Take the second derivative

$$\frac{\partial^2 G(\varphi_1,\tau_0)}{\partial\delta^2} = \sigma(\sigma+1)\frac{\lambda_2^2}{1-\lambda_2}B\left(\varepsilon^\sigma + \mathbb{1}_{\varphi_1,\tau_0,b_L} * (2-\varepsilon)^\sigma - 2\right)\left(\frac{1-\lambda_2\delta}{1-\lambda_2}\right)^{-\sigma-2} \tag{102}$$

As δ increases from 1 to $\delta^-_{\tau_0}$, $\mathbb{1}_{\varphi_1,\tau_0,b_L} = 0$. In this case, φ_1 firms cannot make positive profit under τ_0 conditional on bad belief b_L since τ_0 is not favorable enough. The sign of the first and second derivative depends on the relation between ε^σ and 2. If $\varepsilon^\sigma - 2 > 0$, $\frac{\partial G(\varphi_1,\tau_0)}{\partial\delta} > 0$ and $\frac{\partial^2 G(\varphi_1,\tau_0)}{\partial\delta^2} > 0$, which means $G(\varphi_1,\tau_0)$ is increasing and convex. Otherwise $\frac{\partial G(\varphi_1,\tau_0)}{\partial\delta} < 0$ and $\frac{\partial^2 G(\varphi_1,\tau_0)}{\partial\delta^2} < 0$. Therefore, as $\delta < \delta_{\tau_0}$, the monotonicity and convexity of $G(\varphi_1,\tau_0)$ depends on how good the good belief is. For a high value of good belief, good news can encourage φ_1 firms' early entry while for a low value of good belief, good news can even deter φ_1 firms' early entry. As δ increases from $\delta^-_{\tau_0}$ to $\delta^+_{\tau_0}$, $\mathbb{1}_{\varphi_1,\tau_0,b_L}$ turns from 0 to 1. In this case, exporting under τ_0 conditional on bad belief b_L becomes profitable and there will be an upward jump of the first and second derivative of $G(\varphi_1,\tau_0)$. Both first and second derivative turn to positive since $\varepsilon^\sigma + (2-\varepsilon)^\sigma - 2 > 0$. As δ increases from $\delta^+_{\tau_0}$, $\mathbb{1}_{\varphi_1,\tau_0,b_L} = 1$ and $G(\varphi_1,\tau_0)$ is an increasing and convex function of δ.

K.3 Joint effect of bad and good news

Now we can write the first and second derivative of $Sdiff(\varphi_1,\tau_1)$ regarding to δ. The first derivative is

$$\begin{aligned}\frac{\partial Sdiff(\varphi_1,\tau_1)}{\partial\delta} =& -\mathbb{1}_{\varphi_1,\tau_2,b_H} * \sigma\frac{\lambda_2}{1-\beta}B\varepsilon^\sigma\delta^{-\sigma-1} \\ &+ \sigma\lambda_2 B\left(\varepsilon^\sigma + \mathbb{1}_{\varphi_1,\tau_0,b_L} * (2-\varepsilon)^\sigma - 2\right)\left(\frac{1-\lambda_2\delta}{1-\lambda_2}\right)^{-\sigma-1}\end{aligned} \tag{103}$$

And the second derivative is

$$\begin{aligned}\frac{\partial^2 Sdiff\left(\varphi_1,\tau_1\right)}{\partial\delta^2} =&\mathbb{1}_{\varphi_1,\tau_2,b_H} * \sigma\left(\sigma+1\right)\frac{\lambda_2}{1-\beta}B\varepsilon^{\sigma}\delta^{-\sigma-2} \\ &+\sigma\left(\sigma+1\right)\frac{\lambda_2^2}{1-\lambda_2}B\left(\varepsilon^{\sigma}+\mathbb{1}_{\varphi_1,\tau_0,b_L} * \left(2-\varepsilon\right)^{\sigma}-2\right)\left(\frac{1-\lambda_2\delta}{1-\lambda_2}\right)^{-\sigma-2}\end{aligned} \tag{104}$$

(1) δ **is close to** 1

In this case, we consider a small δ. Since δ is small, $\mathbb{1}_{\varphi_1,\tau_2,b_H}=1$ and $\mathbb{1}_{\varphi_1,\tau_0,b_L}=0$. Exporting under τ_2 conditional on good belief b_H is profitable while exporting under τ_0 conditional on bad belief b_L is not profitable because both τ_2 and τ_0 are close to τ_1. We have

$$\frac{\partial Sdiff\left(\varphi_1,\tau_1\right)}{\partial\delta}|_{\delta\ small} \approx -\sigma\frac{\lambda_2}{1-\beta}B\varepsilon^{\sigma}+\sigma\lambda_2 B\left(\varepsilon^{\sigma}-2\right)<0$$

Conditional on a small δ, as δ increases, for φ_1 firms, the increasing loss under τ_2 dominates the profit change under τ_0 because the loss under τ_2 is for multiple periods while the change under τ_0 is just for one period. For a small δ, $Sdiff\left(\varphi_1,\tau_1\right)<0$ and there are less entrants. If δ increases, the number of entrants will decreases.

(2) **Monotonicity of** $Sdiff\left(\varphi_1,\tau_1\right)$

i) $\varepsilon^{\sigma}-2>0$

In this case, the second derivative of $Sdiff_u(\varphi_1)$ is always positive. In addition, at δ_{τ_0} and δ_{τ_2}, there is an upward jump of the first derivative. As $\delta>\delta_{\tau_0}$ and $\delta>\delta_{\tau_2}$, the first derivative is positive. Therefore, as δ increases from 1 to δ_{max}, there only exist 2 possibilities - either $Sdiff\left(\varphi_1,\tau_1\right)$ decreases always or $Sdiff\left(\varphi_1,\tau_1\right)$ decreases firstly then increases. For a sufficiently large δ, $Sdiff\left(\varphi_1,\tau_1\right)$ can be an increasing function of δ and pass above 0.

ii) $\varepsilon^{\sigma}-2<0$

In this case, as $\delta<\delta_{\tau_0}$, the first derivative of $Sdiff\left(\varphi_1,\tau_1\right)$ is negative and $Sdiff\left(\varphi_1,\tau_1\right)$ decreases in δ. As $\delta>\delta_{\tau_0}$, the second derivative of $Sdiff\left(\varphi_1,\tau_1\right)$ is always positive, which means the first derivative increases in δ and can pass above 0 if δ is sufficiently large. As δ increases from 1 to δ_{max}, there only exist 2 possibilities - either $Sdiff\left(\varphi_1,\tau_1\right)$ decreases always or $Sdiff\left(\varphi_1,\tau_1\right)$ decreases firstly then increases.

L $Sdiff\left(\varphi_1,\tau_1\right)$ as a function of ε

Assuming $f=1$ and substituting b_H by εb_M, b_L by $\left(2-\varepsilon\right)b_M$ and $\varphi_1^{\sigma-1}$ by $\frac{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)f}{\left(b_M^{\sigma}+\frac{1}{2}\frac{\beta}{1-\beta}b_H^{\sigma}\right)k}\tau_1^{\sigma}$, $Sdiff\left(\varphi_1,\tau_1\right)$ is

$$\begin{aligned}&Sdiff\left(\varphi_1,\tau_1\right)= \\ &-\frac{B\varepsilon^{\sigma}-1}{1-\beta}+\frac{\lambda_2}{1-\beta}\max\left\{B\varepsilon^{\sigma}\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}-1,0\right\} \\ &+\left(1-\lambda_2\right)\left(B\varepsilon^{\sigma}\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}+\max\left\{B\left(2-\varepsilon\right)^{\sigma}\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}-1,0\right\}-2B\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}+1\right)\end{aligned} \tag{105}$$

Where $B = \frac{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)}{1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^{\sigma}}$. $B\varepsilon^{\sigma}$ is an increasing function of ε; B is a decreasing function of ε and $B\left(2-\varepsilon\right)^{\sigma}$ is also a decreasing function of ε. Begin with the first and second derivative of $B\varepsilon^{\sigma}$, B and $B\left(2-\varepsilon\right)^{\sigma}$ regarding to ε.

$$\frac{\partial B\varepsilon^{\sigma}}{\partial\varepsilon} = \frac{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)\sigma\varepsilon^{\sigma-1}}{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^{\sigma}\right)^{2}} > 0 \tag{106}$$

$$\frac{\partial B}{\partial\varepsilon} = -\frac{\frac{1}{2}\frac{\beta}{1-\beta}\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)\sigma\varepsilon^{\sigma-1}}{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^{\sigma}\right)^{2}} = -\frac{1}{2}\frac{\beta}{1-\beta}\frac{\partial B\varepsilon^{\sigma}}{\partial\varepsilon} < 0 \tag{107}$$

$$\frac{\partial B\left(2-\varepsilon\right)^{\sigma}}{\partial\varepsilon} = \frac{-\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)\sigma\left(2-\varepsilon\right)^{\sigma-1} - \frac{1}{2}\frac{\beta}{1-\beta}\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)\sigma\varepsilon^{\sigma}\left(2-\varepsilon\right)^{\sigma-1} - \frac{1}{2}\frac{\beta}{1-\beta}\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)\sigma\varepsilon^{\sigma-1}\left(2-\varepsilon\right)^{\sigma}}{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^{\sigma}\right)^{2}} < 0 \tag{108}$$

From (106), (107) and (108), there is a following inequality.

$$\frac{\partial B\varepsilon^{\sigma}}{\partial\varepsilon} + \frac{\partial B\left(2-\varepsilon\right)^{\sigma}}{\partial\varepsilon} - 2\frac{\partial B}{\partial\varepsilon} > 0 \tag{109}$$

From the first order conditions above, we know that $L(\varphi_1,\tau_1)$ is a decreasing function of ε. For φ_1 firms, greater is ε, greater is the net loss under τ_1. $B(\varphi_1,\tau_2)$ is a non decreasing function of ε. For φ_1 firms, greater is ε, (weakly) greater is the gain under τ_2. $G(\varphi_1,\tau_0)$ is an increasing function of ε. For φ_1 firms, greater is ε, greater is the gain under τ_0. Moreover, $L(\varphi_1,\tau_1)+B(\varphi_1,\tau_2)$ is a decreasing function of ε. As ε increases, the increasing loss under τ_1 dominates the (weakly) increasing gain under τ_2. However, for now, it's difficult to tell if $\frac{\partial Sdiff(\varphi_1,\tau_1)}{\partial\varepsilon}$ is positive or not since both increasing and decreasing parts exist. Below, I also give the second derivative.

$$\frac{\partial^2 B\varepsilon^{\sigma}}{\partial\varepsilon^2} = \left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)\sigma\frac{(\sigma-1)\varepsilon^{\sigma-2} - \frac{1}{2}\frac{\beta}{1-\beta}(\sigma+1)\varepsilon^{2\sigma-2}}{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^{\sigma}\right)^{3}} \tag{110}$$

$$\frac{\partial^2 B}{\partial\varepsilon^2} = -\frac{1}{2}\frac{\beta}{1-\beta}\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)\sigma\frac{(\sigma-1)\varepsilon^{\sigma-2} - \frac{1}{2}\frac{\beta}{1-\beta}(\sigma+1)\varepsilon^{2\sigma-2}}{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^{\sigma}\right)^{3}} = -\frac{1}{2}\frac{\beta}{1-\beta}\frac{\partial^2 B\varepsilon^{\sigma}}{\partial^2\varepsilon} \tag{111}$$

$$
\begin{aligned}
&\frac{\partial^2 B\left(2-\varepsilon\right)^\sigma}{\partial \varepsilon^2} = \\
&\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)\sigma\frac{(\sigma-1)\left(2-\varepsilon\right)^{\sigma-2}+\frac{1}{2}\frac{\beta}{1-\beta}(\sigma-1)\varepsilon^\sigma\left(2-\varepsilon\right)^{\sigma-2}+\frac{\beta}{1-\beta}\sigma\varepsilon^{\sigma-1}\left(2-\varepsilon\right)^{\sigma-1}}{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^\sigma\right)^3} \\
&-\frac{1}{2}\frac{\beta}{1-\beta}\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\right)\sigma* \\
&\left[\frac{(\sigma-1)\left(\varepsilon^{\sigma-2}\left(2-\varepsilon\right)^\sigma-\varepsilon^\sigma\left(2-\varepsilon\right)^{\sigma-2}\right)+\frac{1}{2}\frac{\beta}{1-\beta}(\sigma-1)\left(\varepsilon^{2\sigma-2}\left(2-\varepsilon\right)^\sigma-\varepsilon^{2\sigma}\left(2-\varepsilon\right)^{\sigma-2}\right)}{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^\sigma\right)^3}\right. \\
&\left.-\frac{\frac{\beta}{1-\beta}\sigma\left(\varepsilon^{2\sigma-1}\left(2-\varepsilon\right)^{\sigma-1}+\varepsilon^{2\sigma-2}\left(2-\varepsilon\right)^\sigma\right)}{\left(1+\frac{1}{2}\frac{\beta}{1-\beta}\varepsilon^\sigma\right)^3}\right]>0
\end{aligned} \tag{112}
$$

From (110) and (111), it's easy to observe that if $\frac{1}{2}\frac{\beta}{1-\beta}\frac{\sigma+1}{\sigma-1}\varepsilon^\sigma > 1$, $\frac{\partial^2 B\varepsilon^\sigma}{\partial\varepsilon^2} < 0$ and $\frac{\partial^2 B}{\partial\varepsilon^2} > 0$. Otherwise $\frac{\partial^2 B\varepsilon^\sigma}{\partial\varepsilon^2} > 0$ and $\frac{\partial^2 B}{\partial\varepsilon^2} < 0$. We would expect $\frac{1}{2}\frac{\beta}{1-\beta}\frac{\sigma+1}{\sigma-1}\varepsilon^\sigma > 1$ is more likely to be satisfied because normally the discount factor β is not a very small number.[5] However, even if we assume that $\frac{\partial^2 B\varepsilon^\sigma}{\partial\varepsilon^2} < 0$ and $\frac{\partial^2 B}{\partial\varepsilon^2} > 0$, the sign of the second derivative of $Sdiff\left(\varphi_1,\tau_1\right)$ is still ambiguous in some cases. There exist two zero profit cutoffs ε_{τ_2} and ε_{τ_0}. As $\varepsilon > \varepsilon_{\tau_2}$, $\mathbb{1}_{\varphi_1,\tau_2,b_H} = 1$ and $\max\left\{B\varepsilon^\sigma\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}-1,0\right\} = B\varepsilon^\sigma\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}-1>0$. Otherwise $\mathbb{1}_{\varphi_1,\tau_2,b_H}=0$. As $\varepsilon<\varepsilon_{\tau_0}$, $\mathbb{1}_{\varphi_1,\tau_0,b_L}=1$ and $\max\left\{B\left(2-\varepsilon\right)^\sigma\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}-1,0\right\}>0$. Otherwise $\mathbb{1}_{\varphi_1,\tau_0,b_L}=0$. Below I will discuss $Sdiff\left(\varphi_1,\tau_1\right)$ as a function of ε part by part.

L.1 $L(\varphi_1,\tau_1)+B(\varphi_1,\tau_2)$

I start by discussing $L(\varphi_1,\tau_1)$ and $B(\varphi_1,\tau_2)$ together since the sum of these two terms has a good property of monotonicity.

$$
L(\varphi_1,\tau_1)+B(\varphi_1,\tau_2) = -\frac{B\varepsilon^\sigma-1}{1-\beta}+\frac{\lambda_2}{1-\beta}\max\left\{B\varepsilon^\sigma\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}-1,0\right\} \tag{113}
$$

Take the first derivative and we have

$$
\frac{\partial L(\varphi_1,\tau_1)+B(\varphi_1,\tau_2)}{\partial\varepsilon} = \frac{-1+\mathbb{1}_{\varphi_1,\tau_2,b_H}*\lambda_2\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}}{1-\beta}\frac{\partial B\varepsilon^\sigma}{\partial\varepsilon}<0 \tag{114}
$$

Take the second derivative and we have

$$
\frac{\partial^2 L(\varphi_1,\tau_1)+B(\varphi_1,\tau_2)}{\partial\varepsilon^2} = \frac{-1+\mathbb{1}_{\varphi_1,\tau_2,b_H}*\lambda_2\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}}{1-\beta}\frac{\partial^2 B\varepsilon^\sigma}{\partial\varepsilon^2} \tag{115}
$$

As ε increases from 1 to $\varepsilon_{\tau_2}^-$, $L(\varphi_1,\tau_1)+B(\varphi_1,\tau_2) = -\frac{B\varepsilon^\sigma-1}{1-\beta}$ which is a decreasing function of ε. The loss under τ_1 is greater since the value of good belief b_H is greater. As ε increases from $\varepsilon_{\tau_2}^-$

[5] In my model, there is endogenous exit. Therefore, it is only necessary to set a β which is able to capture the exogenous death rate.

to $\varepsilon^+_{\tau_2}$, $\mathbb{1}_{\varphi_1,\tau_2,b_H}$ turns to 1 and $\frac{\lambda_2}{1-\beta}\left(B\varepsilon^\sigma\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}-1\right)$ shows up in (113) whose first derivative is positive. Therefore, there is an upward jump of the first derivative of $L(\varphi_1,\tau_1)+B(\varphi_1,\tau_2)$. In this case, exporting under τ_2 starts to be profitable conditional on good belief. As ε increases from $\varepsilon^+_{\tau_2}$, $L(\varphi_1,\tau_1)+B(\varphi_1,\tau_2)$ is still a decreasing function. However, the loss under τ_1 is partially offset by the gain under τ_2. If we assume $\frac{\partial^2 B\varepsilon^\sigma}{\partial\varepsilon^2}<0$, $\frac{\partial^2 f_{\tau_1}(\varepsilon)+f_{\tau_2}(\varepsilon)}{\partial\varepsilon^2}$ will always be positive and there will be a downward jump at $\varepsilon=\varepsilon_{\tau_2}$.

L.2 $G(\varphi_1,\tau_0)$

We have

$$G(\varphi_1,\tau_0)=(1-\lambda_2)\left(B\varepsilon^\sigma\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}+\max\left\{B\left(2-\varepsilon\right)^\sigma\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}-1,0\right\}-2B\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}+1\right) \quad (116)$$

Take the first derivative and we have

$$\frac{\partial G(\varphi_1,\tau_0)}{\partial\varepsilon}=(1-\lambda_2)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\left(\frac{1}{1-\beta}\frac{\partial B\varepsilon^\sigma}{\partial\varepsilon}+\mathbb{1}_{\varphi_1,\tau_0,b_L}*\frac{\partial B\left(2-\varepsilon\right)^\sigma}{\partial\varepsilon}\right)>0 \quad (117)$$

Take the second derivative

$$\frac{\partial^2 G(\varphi_1,\tau_0)}{\partial\varepsilon^2}=(1-\lambda_2)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\left(\frac{1}{1-\beta}\frac{\partial^2 B\varepsilon^\sigma}{\partial\varepsilon^2}+\mathbb{1}_{\varphi_1,\tau_0,b_L}*\frac{\partial^2 B\left(2-\varepsilon\right)^\sigma}{\partial\varepsilon^2}\right) \quad (118)$$

As ε increases from 1 to $\varepsilon^-_{\tau_0}$, $G(\varphi_1,\tau_0)=(1-\lambda_2)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}B\left(\varepsilon^\sigma+(2-\varepsilon)^\sigma-2\right)$ which is an increasing function of ε. In this case, exporting under τ_0 is still profitable even conditional on bad belief b_L. As ε increases from $\varepsilon^-_{\tau_0}$ to $\varepsilon^+_{\tau_0}$, $\mathbb{1}_{\varphi_1,\tau_0,b_L}$ turns to 0 and $B\left(2-\varepsilon\right)^\sigma\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}-1$ will disappear in (116) whose first derivative is negative. Therefore, there will be an upward jump of the first derivative of $G(\varphi_1,\tau_0)$. In this case, exporting under τ_0 starts to be unprofitable conditional on bad belief. As ε increases from $\varepsilon^+_{\tau_0}$, $G(\varphi_1,\tau_0)$ is an increasing function. If we assume $\frac{\partial^2 B\varepsilon^\sigma}{\partial\varepsilon^2}<0$, $\frac{\partial^2 G(\varphi_1,\tau_0)}{\partial\varepsilon^2}<0$ if $\mathbb{1}_{\varphi_1,\tau_0,b_L}=0$. If $\mathbb{1}_{\varphi_1,\tau_0,b_L}=1$, we cannot tell the sign of the second derivative easily. In addition, for the second derivative, there will be a downward jump at $\varepsilon=\varepsilon_{\tau_0}$.

L.3 Joint effect of ε

The first derivative of $Sdiff\left(\varphi_1,\tau_1\right)$ is

$$\begin{aligned}\frac{\partial Sdiff\left(\varphi_1,\tau_1\right)}{\partial\varepsilon}=&\frac{-1+\mathbb{1}_{\varphi_1,\tau_2,b_H}*\lambda_2\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}}{1-\beta}\frac{\partial B\varepsilon^\sigma}{\partial\varepsilon}\\&+(1-\lambda_2)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\left(\frac{1}{1-\beta}\frac{\partial B\varepsilon^\sigma}{\partial\varepsilon}+\mathbb{1}_{\varphi_1,\tau_0,b_L}*\frac{\partial B\left(2-\varepsilon\right)^\sigma}{\partial\varepsilon}\right)\end{aligned} \quad (119)$$

And the second derivative is

$$\begin{aligned}\frac{\partial^2 Sdiff\left(\varphi_1,\tau_1\right)}{\partial\varepsilon^2}=&\frac{-1+\mathbb{1}_{\varphi_1,\tau_2,b_H}*\lambda_2\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma}}{1-\beta}\frac{\partial^2 B\varepsilon^\sigma}{\partial\varepsilon^2}\\&+(1-\lambda_2)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\left(\frac{1}{1-\beta}\frac{\partial^2 B\varepsilon^\sigma}{\partial\varepsilon^2}+\mathbb{1}_{\varphi_1,\tau_0,b_L}*\frac{\partial^2 B\left(2-\varepsilon\right)^\sigma}{\partial\varepsilon^2}\right)\end{aligned} \quad (120)$$

We can easily observe that as $\varepsilon \to 1$, $Sdiff\left(\varphi_1, \tau_1\right) \to 0$ and $\varphi_{1u} \to \varphi_1$. Moreover, in this case, since there is no learning, the entry threshold is just the zero profit cutoff conditional on τ_1 and b_M.

(1) ε **is close to** 1

As ε is close to 1, $\mathbb{1}_{\varphi_1,\tau_2,b_H} = 0$ and $\mathbb{1}_{\varphi_1,\tau_0,b_L} = 1$. In this case, good belief is not too good and bad belief is not too bad. Therefore, exporting under τ_2 conditional on good belief is not profitable and exporting under τ_0 conditional on bad belief is profitable. The first derivative is

$$\frac{\partial Sdiff\left(\varphi_1, \tau_1\right)}{\partial \varepsilon}|_{\varepsilon\, small} = -\frac{1}{1-\beta}\frac{\partial B\varepsilon^{\sigma}}{\partial \varepsilon} + (1-\lambda_2)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}\left(\frac{1}{1-\beta}\frac{\partial B\varepsilon^{\sigma}}{\partial \varepsilon} + \frac{\partial B\left(2-\varepsilon\right)^{\sigma}}{\partial \varepsilon}\right) < 0 \tag{121}$$

As ε is small, $\frac{\partial B\varepsilon^{\sigma}}{\partial \varepsilon} \approx \frac{\sigma}{1+\frac{1}{2}\frac{\beta}{1-\beta}}$ and $\frac{1}{1-\beta}\frac{\partial B\varepsilon^{\sigma}}{\partial \varepsilon} + \frac{\partial B(2-\varepsilon)^{\sigma}}{\partial \varepsilon} \approx 0$. In this case, under τ_0, as ε increases, the extra benefit from good belief will be offset by the extra loss from bad belief. Therefore, the loss under τ_1 dominates and we have $\frac{\partial Sdiff_u(\varphi_1)}{\partial \varepsilon}|_{\varepsilon\, small} < 0$. I don't make a further discussion on the second order condition here since it is not obvious.[6]

(2) $\varepsilon > \varepsilon_{\tau_2}$ **and** $\varepsilon > \varepsilon_{\tau_0}$

In this case, $\mathbb{1}_{\varphi_1,\tau_2,b_H} = 1$ and $\mathbb{1}_{\varphi_1,\tau_0,b_L} = 0$. The first derivative is

$$\frac{\partial Sdiff\left(\varphi_1, \tau_1\right)}{\partial \varepsilon}|_{\varepsilon\, large} = \frac{-1+\lambda_2\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma} + (1-\lambda_2)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}}{1-\beta}\frac{\partial B\varepsilon^{\sigma}}{\partial \varepsilon} \tag{122}$$

And the second derivative is

$$\frac{\partial^2 Sdiff\left(\varphi_1, \tau_1\right)}{\partial \varepsilon^2}|_{\varepsilon\, large} = \frac{-1+\lambda_2\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma} + (1-\lambda_2)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}}{1-\beta}\frac{\partial^2 B\varepsilon^{\sigma}}{\partial \varepsilon^2} \tag{123}$$

As $\tau_2 = \tau_0 = \tau_1$, $-1+\lambda_2\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma} + (1-\lambda_2)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma} = 0$. The sign of $\frac{\partial Sdiff(\varphi_1,\tau_1)}{\partial \varepsilon}|_{\varepsilon\, large}$ depends on the sign of $-1+\lambda_2\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma} + (1-\lambda_2)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma}$, which is similar to the result in sunk cost learning model. In this case, exporting under τ_2 conditional on good belief is profitable and exporting under τ_0 conditional on bad belief is unprofitable. The former term is an increasing function of ε and the later term is a decreasing function of ε. As $\varepsilon > \varepsilon_{\tau_2}$ and $\varepsilon > \varepsilon_{\tau_0}$, we introduce an increasing term and get rid of a decreasing term. Therefore, if $-1+\lambda_2\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma} + (1-\lambda_2)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma} < 0$, $\frac{\partial Sdiff(\varphi_1,\tau_1)}{\partial \varepsilon}$ will always be negative. However, if $-1+\lambda_2\left(\frac{\tau_2}{\tau_1}\right)^{-\sigma} + (1-\lambda_2)\left(\frac{\tau_0}{\tau_1}\right)^{-\sigma} > 0$, as $\varepsilon < \varepsilon_{\tau_2}$ or $\varepsilon < \varepsilon_{\tau_0}$, the sign of the first derivative is not clear and we cannot tell the sign of the second derivative easily either. Therefore, we don't have a clear conclusion in this case.

M Data set construction

I build indicator of positive bilateral trade flow at product level between 1995 and 2019 using BACI database, which measures the probability of HS6 product entry in foreign market. HS6 code is converted to 1992 version. The detailed steps are as follows:

[6] It's more likely to be positive.

1. The HS6 MFN tariff data is from the WITS Trains database. In the raw data, importers are all individual countries I can choose and exporters are "world" since I use MFN and BND tariff (invariant across WTO exporters). The countries I consider are WTO members. Both BND and MFN are simple average at HS6 level. I only keep HS6 tariff lines whose MFN=BND>0.

2. I only keep the US and Canada as importers. I believe these 2 countries have more complete trade flows records in BACI data. And I don't want to introduce too many observations. Using "joinby" and gravity database, I assign each importer-hs6-year all possible potential WTO exporters. Also, the Gravity database provides yearly RTA information (0 for no RTA agreement and 1 for RTA agreement in a given year).[7]

3. In BACI database, I only keep the US and CAN as importers. Then I merge BACI data with the above tariff-gravity dataset. For observations in tariff-gravity dataset that cannot merge with BACI observations, I assign 0 as trade value. Then I can build a positive trade flow indicator at ex-hs6-des-year level (1 if positive and 0 if 0). There are 149 WTO countries included. 1,207,837 out of 6,271,928 observations have positive trade flow. The graph below shows RTA enforcement during 1995-2019 given the US & CAN as importer.

[7] Gravity data on Regional Trade Agreements (RTAs) (rta, rta_type, rta_coverage) is taken from the WTO's (2020) "Regional Trade Agreements Information System (RTA-IS)". For each RTA, this dataset lists the RTA name, the coverage (whether it's goods, services or both), the type of RTA, the date of entry into force for the part on goods and for the part on services (the two may differ), the original signatories, and specific entry or exit dates for additional signatories. A country pair is considered as being in a RTA in a given year as soon as the RTA was in force at least one day during this year. For now, I dismiss the coverage and the type of RTA.

```
/*list iso3_o iso3_d year if post1==1
       +--------------------------+
       | iso3_o   iso3_d    year |
       |--------------------------|
  302. |    AUS      CAN    2018 |
  314. |    AUS      USA    2005 |
  585. |    BHR      USA    2006 |
  822. |    BRN      CAN    2018 |
  988. |    CHE      CAN    2009 |
       |--------------------------|
 1026. |    CHL      CAN    1997 |
 1058. |    CHL      USA    2004 |
 1320. |    COL      CAN    2011 |
 1346. |    COL      USA    2012 |
 1385. |    CRI      CAN    2002 |
       |--------------------------|
 1414. |    CRI      USA    2006 |
 1650. |    DOM      USA    2006 |
 1794. |     EU      CAN    2017 |
 2242. |    GTM      USA    2006 |
 2375. |    HND      CAN    2014 |
       |--------------------------|
 2392. |    HND      USA    2006 |
 2612. |    ISL      CAN    2009 |
 2650. |    ISR      CAN    1997 |
 2760. |    JOR      CAN    2012 |
 2769. |    JOR      USA    2001 |
       |--------------------------|
 2811. |    JPN      CAN    2018 |
 3042. |    KOR      CAN    2015 |
 3064. |    KOR      USA    2012 |
 3208. |    LIE      CAN    2009 |
 3446. |    MAR      USA    2006 |
       |--------------------------|
 4085. |    MYS      CAN    2018 |
 4296. |    NIC      USA    2006 |
 4324. |    NOR      CAN    2009 |
 4415. |    NZL      CAN    2018 |
 4471. |    OMN      USA    2009 |
       |--------------------------|
 4548. |    PAN      CAN    2013 |
 4570. |    PAN      USA    2012 |
 4592. |    PER      CAN    2009 |
 4617. |    PER      USA    2009 |
 5033. |    SGP      CAN    2018 |
       |--------------------------|
 5044. |    SGP      USA    2004 |
 5194. |    SLV      USA    2006 |
 5837. |    UKR      CAN    2017 |
 6039. |    VNM      CAN    2019 |
       +--------------------------+
*/
```

lnpop_o	ln Population of origin, total in million
lnpop_d	ln Population of destination, total in million
lngdpcap_o	ln Gross Domestic Product per capita of origin (current US)
lngdpcap_d	ln Gross Domestic Product per capita of destination (current US)
lndistw	Weighted bilateral distance between origin and destination in kilometer (population weighted)
Contiguity	1 for contiguity
Common language	1 for common official or primary language
Common legal origins before transition	1 if common legal origins before transition
Common legal origins after transition	1 if common legal origins after transition

N Other regression results

i is export country and j is destination. mfn_{jpt} is simple average MFN tariff at HS6 level and p denotes HS6 product. RTA_{ijt} is Regional Trade Agreement dummy. $RTA_{ijt} = 1$ if RTA is enforced at ijt level. $positive_{ijpt} = 1$ if positive trade flow at $ijpt$ level. $pre1_{ijt} = 1$ for one year before RTA enforcement and $pre1_{ijt} = 0$ for all the other years. $post1_{ijt} = 1$ for the first year of RTA enforcement and $post1_{ijt} = 0$ for all the other years. The same rule is applied to $pre2_{ijt} = 1$, $pre3_{ijt} = 1$, $post2_{ijt} = 1$, $post3_{ijt} = 0$. Gravity controls are $lnpop$, $lngdpcap$, $lndistw$, $Contiguity$, $common\,language$ and $common\,legal\,origins\,before\,/\,after\,transition$. it is export country-year dummy. jt is destination-year dummy. ij is export country-destination dummy. $jhs2t$ is destination-HS2 sector-year dummy.

N.1 Product-level entry

$$positive_{ijpt} = \beta_1 * mfn_{jpt} + \beta_2 * pre1_{ijt} + \beta_3 * RTA_{ijt} + \beta_g * gravity + \beta_d * Dummy + \epsilon_{ijpt} \quad (124)$$

Tab. 5: RTA and one period pre-rta

	(1) positive	(2) positive	(3) positive	(4) positive	(5) positive
mfn	0.000434***	0.000453***	0.000368***	0.000375***	-0.000428***
	(2.93)	(4.59)	(4.32)	(4.45)	(-6.36)
pre1	0.189***	0.0726***	0.00735	0.0240***	0.0240***
	(4.68)	(2.65)	(0.48)	(3.52)	(3.52)
RTA	0.270***	0.0521**	-0.0222	-0.00876	-0.00876
	(4.89)	(1.97)	(-1.02)	(-0.88)	(-0.88)
lnpop_o		0.0879***			
		(12.67)			
lnpop_d		0.0210**			
		(2.48)			
lngdpcap_o		0.0881***			
		(11.20)			
lngdpcap_d		-0.0814***			
		(-4.47)			
lndistw		-0.0618***	-0.315***		
		(-2.92)	(-5.83)		
Contiguity		0.197***	0.0183		
		(4.30)	(0.96)		
common language		0.0474*	0.0210*		
		(1.94)	(1.83)		
Common legal origins before transition		0.133***	-0.171***		
		(3.03)	(-4.12)		
Common legal origins after transition		-0.0969***	0.200***		
		(-3.04)	(4.86)		
fixed effect			it, jt	it, jt, ij	it, jhs2t, ij
cluster	ij	ij	ij	ij	ij
Observations	5966527	5818513	5856185	5966527	5966527

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

$$D.positive_{ijpt} = \beta_1 * D.mfn_{jpt} + \beta_2 * D.pre1_{ijt} + \beta_3 * D.RTA_{ijt} + \beta_g * D.gravity + \beta_d * Dummy + \epsilon_{ijpt} \tag{125}$$

Tab. 6: RTA and one period pre-rta first difference

	(1) D.positive	(2) D.positive	(3) D.positive	(4) D.positive
D.mfn	0 (.)	0 (.)	0 (.)	0 (.)
D.pre1	0.00979*** (2.96)	0.00981*** (3.02)	0.0172*** (3.35)	0.0172*** (3.35)
RTA	0.00861** (2.19)	0.00893** (2.32)	0.00694 (1.33)	0.00695 (1.33)
D.lnpop_o		-0.00106 (-0.06)		
D.lnpop_d		0.00754 (0.39)		
D.lngdpcap_o		-0.00558* (-1.94)		
D.lngdpcap_d		0.0286*** (4.94)		
fixed effect			it, jt	it, jhs2t
cluster	ij	ij	ij	ij
Observations	5265138	5219306	5265138	5265138

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

$$positive_{ijpt} = \beta_1 * mfn_{jpt} + \beta_2 * pre1_{ijt} + \beta_3 * pre2_{ijt} + \beta_4 * pre3_{ijt} + \beta_5 * post1_{ijt} + \beta_6 * post2_{ijt} + \beta_7 * post3_{ijt} + \beta_g * gravity + \beta_d * Dummy + \epsilon_{ijpt} \quad (126)$$

Tab. 7: Dynamic effect - 3 pre and 3 post

	(1) positive	(2) positive	(3) positive	(4) positive	(5) positive
mfn	0.000463***	0.000458***	0.000380***	0.000387***	-0.000431***
	(3.01)	(4.54)	(4.31)	(4.44)	(-6.27)
pre1	0.165***	0.0662***	0.00806	0.0296***	0.0296***
	(4.13)	(2.70)	(0.55)	(3.90)	(3.90)
pre2	0.158***	0.0655**	-0.0225**	0.0119*	0.0119*
	(3.91)	(2.59)	(-2.21)	(1.91)	(1.91)
pre3	0.161***	0.0636**	-0.00537	0.0196***	0.0196***
	(3.73)	(2.41)	(-0.46)	(3.07)	(3.07)
post1	0.166***	0.0622**	-0.00677	0.00959	0.00958
	(4.22)	(2.56)	(-0.50)	(1.44)	(1.44)
post2	0.172***	0.0571**	-0.00254	0.0100	0.0100
	(4.31)	(2.34)	(-0.22)	(1.55)	(1.56)
post3	0.157***	0.0446*	-0.00572	0.00218	0.00220
	(3.76)	(1.67)	(-0.42)	(0.27)	(0.27)
lnpop_o		0.0890***			
		(13.24)			
lnpop_d		0.0206**			
		(2.47)			
lngdpcap_o		0.0890***			
		(11.85)			
lngdpcap_d		-0.0555***			
		(-2.81)			
lndistw		-0.0665***	-0.307***		
		(-3.18)	(-5.60)		
Contiguity		0.232***	0.0195		
		(4.43)	(1.00)		
Common language		0.0488**	0.0216*		
		(1.97)	(1.87)		
Common legal origins before transition		0.143***	-0.129***		
		(3.38)	(-3.81)		
Common legal origins after transition		-0.107***	0.159***		
		(-3.65)	(4.75)		
fixed effect			it, jt	it, jt, ij	it, jhs2t, ij
cluster	ij	ij	ij	ij	ij
Observations	5543883	5408010	5442435	5543883	5543883

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

N.2 Country-pair-level entry share

$$positive\_share_{ijt} = \frac{Number\,of\,products\,whose\,mfn = bnd\,with\,positive\,trade\,flow\,within\,ijt}{Number\,of\,products\,whose\,mfn = bnd\,within\,jt} \tag{127}$$

The denominator gives number of products for which country j has set MFN=BND>0 at t. The nominator gives number of mfn=bnd products with positive trade flow between i and j at t. m_mfn_{jt} is just simple average of MFN of the products whose mfn=bnd within jt.

$$positive\_share_{ijt} = \beta_1 * m\_mfn_{jt} + \beta_2 * pre1_{ijt} + \beta_3 * RTA_{ijt} + \beta_g * gravity_{it/jt} + \beta_d * Dummy + \epsilon_{ijt} \tag{128}$$

Tab. 8: RTA and one period pre-rta

	(1) positive_share	(2) positive_share	(3) positive_share	(4) positive_share
m_mfn	0.00863^{**}	0.0179^{***}	0.0121^{***}	-0.463^{***}
	(2.30)	(4.30)	(4.79)	(-2.62)
pre1	0.214^{***}	0.0827^{***}	0.0597^{***}	0.0256^{**}
	(5.97)	(3.57)	(4.98)	(2.43)
RTA	0.295^{***}	0.0639^{***}	0.0628^{***}	-0.0161
	(5.62)	(2.77)	(4.13)	(-1.05)
lnpop_o		0.0808^{***}	-0.0455	
		(14.80)	(-1.20)	
lnpop_d		0.0502^{***}	-0.0490	
		(4.85)	(-0.83)	
lngdpcap_o		0.0799^{***}	0.00690	
		(12.47)	(0.71)	
lngdpcap_d		-0.0953^{***}	0.0506^{***}	
		(-3.06)	(3.42)	
lndistw		-0.0453^{***}		
		(-2.73)		
Contiguity		0.226^{***}		
		(4.69)		
Common language		0.0370^{**}		
		(2.10)		
Common legal origins before transition		0.107^{**}		
		(2.40)		
Common legal origins after transition		-0.0713^{*}		
		(-1.84)		
fixed effect			t, ij	it, jt, ij
cluster	ij	ij	ij	ij
Observations	5994	5837	5943	5994

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

$$D.positive_{ijt} = \beta_1 * D.m\_mfn_{jt} + \beta_2 * D.pre1_{ijt} + \beta_3 * D.RTA_{ijt} + \beta_g * D.gravity_{it/jt} + \beta_d * Dummy + \epsilon_{ijt} \quad (129)$$

Tab. 9: RTA and one period pre-rta first difference

	(1) D.positive_share	(2) D.positive_share	(3) D.positive_share
D.m_mfn	0.000763	-0.000122	-0.291***
	(1.09)	(-0.19)	(-2.77)
D.pre1	0.0125**	0.0122**	0.0237***
	(2.36)	(2.38)	(3.54)
RTA	0.00466	0.00696	0.00420
	(0.65)	(1.02)	(0.46)
D.lnpop_o		-0.0241	
		(-1.42)	
D.lnpop_d		0.0382	
		(1.25)	
D.lngdpcap_o		-0.0112***	
		(-3.62)	
D.lngdpcap_d		0.0996***	
		(11.13)	
fixed effect			it, jt
cluster	ij	ij	ij
Observations	5698	5635	5698

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

$$\begin{aligned} positive\_share_{ijt} = & \beta_1 * m\_mfn_{jt} + \beta_2 * pre1_{ijt} + \beta_3 * pre2_{ijt} + \beta_4 * pre3_{ijt} + \beta_5 * post1_{ijt} \\ & + \beta_6 * post2_{ijt} + \beta_7 * post3_{ijt} + \beta_g * gravity_{it/jt} + \beta_d * Dummy + \epsilon_{ijt} \end{aligned} \quad (130)$$

Tab. 10: Dynamic effect - 3 pre and 3 post

	(1) positive_share	(2) positive_share	(3) positive_share	(4) positive_share
m_mfn	0.0120***	0.0188***	0.0107***	-0.417**
	(2.94)	(4.51)	(4.25)	(-2.08)
pre1	0.190***	0.0720***	0.0364***	0.0282***
	(5.29)	(3.38)	(3.91)	(2.85)
pre2	0.179***	0.0626***	0.0221**	0.0106
	(4.85)	(2.75)	(2.28)	(1.27)
pre3	0.184***	0.0728***	0.0234***	0.0265***
	(4.95)	(3.37)	(2.73)	(3.02)
post1	0.187***	0.0643***	0.0262**	0.00664
	(5.12)	(2.89)	(2.35)	(0.84)
post2	0.189***	0.0605***	0.0327***	0.00760
	(5.23)	(2.98)	(2.92)	(0.97)
post3	0.175***	0.0525**	0.0302**	-0.00449
	(4.45)	(2.30)	(2.13)	(-0.37)
lnpop_o		0.0815***	-0.0561	
		(15.32)	(-1.39)	
lnpop_d		0.0518***	-0.0882**	
		(5.29)	(-2.00)	
lngdpcap_o		0.0810***	0.00375	
		(13.28)	(0.37)	
lngdpcap_d		-0.0807***	0.0641***	
		(-2.68)	(4.24)	
lndistw		-0.0493***		
		(-2.92)		
Contiguity		0.266***		
		(5.14)		
Common language		0.0399**		
		(2.26)		
Common legal origins before transition		0.121***		
		(2.94)		
Common legal origins after transition		-0.0865**		
		(-2.53)		
fixed effect			t, ij	it, jt, ij
cluster	ij	ij	ij	ij
Observations	5475	5332	5428	5475

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Below I use two-way fixed effects estimation with heterogeneous treatment effects. The lth placebo compares first-time switchers' and not-yet switchers' outcome evolution, from the l+1th to the lth period before first-time switchers' treatment changes. Thus, the lth placebo assesses if parallel trends holds over 2 consecutive periods, l periods before switchers switch. dynamic() gives the number of dynamic treatment effects to be estimated. DID_l, the estimator of the lth dynamic effect, compares first-time switchers' and not-yet switchers' outcome evolution, from the last period before first-time switchers' treatment changes to the lth period after that change. CI is 95% confidence interval.

1. Control: m_mfn

```
. did_multiplegt positive_share ij t rta , robust_dynamic dynamic(2)  placebo(
> 3) jointtestplacebo controls(m_mfn)  breps(2000) cluster(ij) covariances sav
> e_results(W:\empirical test\baci_test_92\US_CAN_importer\2way_share_mfn_USCA
> N.dta)

DID estimators of the instantaneous treatment effect, of dynamic treatment
effects if the dynamic option is used, and of placebo tests of the parallel
trends assumption if the placebo option is used. The estimators are robust
to heterogeneous effects, and to dynamic effects if the robust_dynamic
option is used.
```

	Estimate	SE	LB CI	UB CI	N
Effect_0	-.0067846	.0043388	-.0152886	.0017193	3453
Effect_1	-.0010882	.0056016	-.0120673	.0098909	3196
Effect_2	-.0035551	.0089346	-.021067	.0139568	2921
Placebo_1	.0150412	.0055408	.0041812	.0259013	3142
Placebo_2	-.0023299	.0042228	-.0106065	.0059468	2912
Placebo_3	.0071393	.0036183	.0000475	.0142311	2876

	Switchers
Effect_0	39
Effect_1	38
Effect_2	32
Placebo_1	38
Placebo_2	36
Placebo_3	36

2. Controls: m_mfn and gravity controls

```
. did_multiplegt positive_share ij t rta , robust_dynamic dynamic(2)  placebo(
> 3)  jointtestplacebo controls(m_mfn lnpop_o lnpop_d lngdpcap_o lngdpcap_d)
> breps(2000) cluster(ij) covariances  save_results(W:\empirical test\baci_tes
> t_92\US_CAN_importer\2way_share_mfn_gravity_USCAN.dta)
```

DID estimators of the instantaneous treatment effect, of dynamic treatment effects if the dynamic option is used, and of placebo tests of the parallel trends assumption if the placebo option is used. The estimators are robust to heterogeneous effects, and to dynamic effects if the robust_dynamic option is used.

	Estimate	SE	LB CI	UB CI	N
Effect_0	-.0049187	.0042199	-.0131897	.0033523	3418
Effect_1	.0019626	.0052048	-.0082388	.012164	3150
Effect_2	.0005389	.0085697	-.0162576	.0173355	2867
Placebo_1	.0159139	.0051407	.0058381	.0259897	3096
Placebo_2	-.0008452	.0040921	-.0088658	.0071754	2858
Placebo_3	.0089613	.003686	.0017367	.0161859	2822

	Switchers
Effect_0	39
Effect_1	38
Effect_2	32
Placebo_1	38
Placebo_2	36
Placebo_3	36