Dynamic Decisions on Import: Evidence from Chinese Firm-Level Data

Kensuke Suzuki*

April 29, 2019

1 Motivation

In the literature of international trade and development economics, a number of studies have investigated the impact of using imported intermediate inputs on firm's performance. In the context of developing countries, in particular, using imported inputs is an important channel through which a firm can increase its productivity. As Kasahara and Rodrigue (2008) argues, importing firms may use more input varieties than non-importers as there are varieties only available in the import market but not in the domestic market. Provided the imperfect substitution across intermediate inputs, this will lead to the reduction in marginal cost of production *à la* Ethier (1982). Halpern et al. (2015) address the quality dimension of imported intermediate input and argue that using imported inputs allows a firm in developing countries to take advantage of foreign advanced technology embodied in the intermediate goods. Therefore, importing may allow for firms to save input expenditure conditional on quality. Wile these benefits are delivered to firms immediately upon importing, previous works also found the evidence of dynamic productivity implication; import experience will positively affect the firm's production efficiency in the future due to, for example, interaction with sophisticated upstream sellers of inputs. This is colloquially referred to learning-by-importing effect.

*PhD student, Department of Economics, the Pennsylvania State University. This study is undertaken as an empirical methods project in ECON512 Empirical Methods (Professor Paul Grieco) and a mini-research project in ECON570 Development Economics (Professor Kala Krishna). The China Industrial Firm Statistics Dataset and China Custom Export-Import Dataset are provided by Professor Hong Ma at Tsinghua University.

The other strand of literature studying firm's trading pattern has found that importers are likely to be larger and more productive than non-importers (Bernard et al., 2012). In the standard Melitz (2003) type of firm heterogeneity model, productivity sorting of firms into importing can be explained by the existence of fixed and sunk costs of import, i.e., firm needs to incur the sunk and fixed cost to initiate and keep foreign sourcing. Only firms that are sufficiently productive find it profitable to import intermediate goods. In the data, however, the productivity sorting of firms into importers and non-importers is not perfect. Looking at Chinese firm-level data, productivity distributions (e.g., value added per worker) of importers and non-importers are substantially overlapping. In other words, there are substantial mass of high productive, but not importing firms, and *vise versa*. This imperfect sorting may be attributed to the heterogeneity of fixed and sunk cost of import.[1]

Given these emphasis, this paper develops and estimates the dynamic structural model of firm's import decision. Static component of the model is based on Kasahara and Rodrigue (2008). The model features two benefits from using the imported inputs. The first one is the static cost reduction effect. Importers may access to more varieties in the foreign markets compared to non-importers which are restricted to source everything from domestic market. This will reduce the marginal cost due to producer analogue of love-of-variety effect. The second benefit from using imported input is dynamic learning-by-importing effect. The model also features the heterogeneous startup sunk cost and fixed cost by introducing the stochastic formulation as in Aw et al. (2011) and Bai et al. (2017). I assume that each firm draws the relevant import cost each period from the known distribution which depends on the previous period import status. Given the learning-by-importing effect and existence of sunk and fixed cost of import, the firm's import decision is dynamic.

I estimate the model using the Chinese firm-level data which covers the period 2000–2006. The estimation of the model can be decomposed into two stages. In the first stage, I estimate the firm's revenue production function and the Markov process of productivity evolution which identify the static and dynamic productivity implications of import. Since the productivity is not observed in the data, I employ Levinsohn and Petrin's (2003) control function technique to control for the unobserved productivity. The computational technique used in this first stage is ordinary least squares and nonlinear least squares.

In the second stage, I estimate the dynamic parameters, i.e., distribution parameters of fixed

[1]The other rationale would be the vertical integration. It may be the case that more productive firms are able to integrate their production process vertically so that they do not have to source intermediate inputs from abroad. This is beyond the scope of this study.

and sunk cost. I employ the nested fixed point algorithm *à la* Rust (1987). This involves solving a set of firm's value functions in the inner loop, which is used to obtain the conditional choice probability of binary import decision. Since the firm's state space contains continuous variables, namely, capital stock and productivity, I discretized the state space by employing Rust's random grid approximation. Conditional choice probability is then used to construct the likelihood of the data. Dynamic parameters are estimated by directly maximizing the likelihood. In addition to the maximum likelihood estimators, I also evaluate the likelihood by employing the Bayesian Markov Chain Monte Carlo (MCMC) estimator. I compare the results from the two methods by assessing the fit of the aggregate moment, i.e., import participation rate over time implied by the model and the data. Finally, I implemented a simple counterfactual analysis, the government's subsidy on sunk and/or fixed cost of import. I examine the trajectory of import participation rates under the different schemes of import subsidy.

The rest of this paper is structured as follows: Section 2 outlines the model, Section 3 explains the empirical specification of the model and estimation strategy, Section 4 demonstrates the empirical results, Section 5 implements counterfactual analysis, and Section 6 concludes.

2 The Model

2.1 Demand side

I focus on a firm's import decision in a particular industry. Both the demand and supply sides are modeled for a given industry. For notational simplicity, I suppress industry index. The utility function defined over the varieties is the CES utility function. Utility of representative consumer in period t is given by

$$U_t = U\left(\left\{Q_{it}\right\}_{i\in\Omega}\right) = \left[\int_{i\in\Omega} \left(Q_{it}\right)^{\frac{\sigma-1}{\sigma}} di\right]^{\frac{\sigma}{\sigma-1}}, \tag{1}$$

where Q_{it} denotes consumption of variety produced by firm $i \in \Omega$ and Ω is a set of varieties (firms). $\sigma > 1$ is elasticity of substitution. In the current manuscript, I do not distinguish domestic market from the foreign markets so that there is a single market in the economy. Given the Dixit-Stiglitz monopolistic competition, firm-level demand can be written as,

$$Q_{it} = \left(\frac{P_{it}}{P_t}\right)^{-\sigma} \frac{E_t}{P_t} = \Phi_t (P_{it})^{-\sigma}, \tag{2}$$

where E_t is total expenditure for the industry and P_t is the CES price index defined by,

$$P_t = \left[\int_{i\in\Omega} (P_{it})^{1-\sigma}\, di\right]^{\frac{1}{1-\sigma}},$$

P_{it} is price of each variety. $\Phi_t \equiv E_t/P_t^{1-\sigma}$ captures industry aggregate variables.

2.2 Production function and cost reduction effect from importing

I follow Kasahara and Rodrigue (2008) for the specification of production function. Output of firm i at time t is given by

$$Y_{it} = \exp(\omega_{it}) K_{it}^{\beta_k} L_{it}^{1-\beta_x} \left[\int_0^{N_i(d_{it})} X(j)^{\frac{\theta-1}{\theta}}\, dj\right]^{\frac{\theta}{\theta-1}\beta_x} \tag{3}$$

where $\exp(\omega_{it})$ is Hicks-neutral productivity, K_{it} is capital input which is assumed to be the fixed factor, and L_{it} is labor input. The firms uses the CES aggregate of input varieties $X(j)$ with elasticity $\theta > 1$ where $j \in [0, N_i(d_{it})]$. $N_i(d_{it})$ is the measure of intermediate input which depends on importing status d_{it}. $d_{it} = 1$ if a firm i is importer in period t and 0 otherwise. This specification implies that importer and non-importer use different sets of input varieties. Since importer may access to intermediate inputs which are available only abroad but not in the domestic market, $N_i(1)$ would be larger than $N_i(0)$. Given the imperfect substitution across intermediate inputs, the more intermediate varieties a firm uses, the more the firm can produce *ceteris paribus* (Ethier, 1982). This is the source of static benefits of using imported intermediate input in my theoretical framework. I will come back to this point shortly. I assume that production function is constant return to scale in labor input and intermediate input.

As in Kasahara and Rodrigue, consider the equilibrium where all intermediate goods are symmetrically produced at the level of $X(j) = \overline{X}$. Then I can rewrite equation (3) as

$$Y_{it} = \exp(\omega_{it}) N_i(d_{it})^{\frac{1}{\theta-1}\beta_x} K_{it}^{\beta_k} L_{it}^{1-\beta_x} X_{it}^{\beta_x} \tag{4}$$

where $X_{it} = N_i(d_{it})\overline{X}$ is total intermediate input expenditure.

For given factor prices and price index for intermediate input, firm's cost minimization problem gives the marginal cost function as follows:

$$MC_{it} = \exp(-\omega_{it}) B_0 w_t^{1-\beta_x} p_{it}^{\beta_x} K_{it}^{\beta_k} N_i(d_{it})^{-\frac{\beta_x}{\theta-1}} \tag{5}$$

where w_t and p_t, respectively, are wage and intermediate input price index, and $B_0 \equiv (1-\beta_x)^{-(1-\beta_x)}\beta_x^{-\beta_x}$ is constant. It is important to note that with imperfect substitution across intermediate inputs, i.e., $-\beta_x/(\theta-1) < 0$, marginal cost is decreasing in the measure of intermediate input $N_i(d_{it})$. This is producer analogue of love-of-variety effect. If a importer uses more varieties than non-importer, i.e., $N_i(1) > N_i(0)$, importer's marginal cost is lower for other things being equal. This captures the static cost reduction effect from using imported inputs.

There are two challenges in taking the model to the data. First, the factor price and input price index are not observable in the data. Therefore, a time dummy D_t captures them in the empirical specification. The other challange is that the exact value of $N_i(d_{it})$ is not available. However, under the assumption of symmetrically produced intermediate inputs, I can calculate the ratio of total intermediate inputs *vis-a-vis* domestically produced intermediate inputs. I will proxy $N_i(d_{it})$ by N_{it}:

$$N_{it} \equiv \frac{X_{it}}{X_{it}^D} = \frac{X_{it}}{X_{it} - X_{it}^I} = \frac{N_i(d_{it}=1)\overline{x}}{N_i(d_{it}=0)\overline{x}} = \frac{N_i(d_{it}=1)}{N_i(d_{it}=0)} \tag{6}$$

X_{it} is the total intermediate input, X_{it}^D is the domestic intermediate input, and X_{it}^I is the total imported intermediate input. By construction, N_{it} takes unity if a firms is non-importer ($d_{it} = 0$) and greater than one otherwise. N_{it} can be seen as the intensive margin of import. Econometric specification for the log marginal cost is:

$$mc_{it} = \beta_0 + \beta_t D_t + \beta_k k_{it} + \beta_n n_{it} - \omega_{it} \tag{7}$$

where lower case character indicates the log of the corresponding upper case variable. β_n captures static benefit from import as discussed above.

Given the CES demand system and monopolistic competition, each firm prices at constant markup over marginal cost:

$$P_{it} = \frac{\sigma}{\sigma - 1} MC_{it}, \tag{8}$$

By combining firm's pricing rule (8) and demand function (2) we can compute the revenue as:

$$\begin{aligned} R_{it} &= P_{it} Q_{it} \\ &= \left(\frac{\sigma}{\sigma - 1} \right)^{1-\sigma} (MC_{it})^{1-\sigma} \Phi_t \\ &= R_{it}(\Phi_t, k_{it}, n_{it}, \omega_{it}) \end{aligned} \tag{9}$$

Taking log and substitute $\ln(MC_{it})$ with the expression obtained in equation (7) gives

$$r_{it} = (1 - \sigma) \ln \left(\frac{\sigma}{\sigma - 1} \right) + \ln(\Phi_t) + (1 - \sigma) \Big\{ \beta_0 + \beta_t D_t + \beta_k k_{it} + \beta_n n_{it} - \omega_{it} \Big\} \tag{10}$$

Given the assumption on the Dixit and Stiglitz form of consumer preferences and monopolistic competition, firm's operation profit Π_{it} is constant share of revenue:

$$\Pi_{it} = \frac{1}{\sigma} R_{it}(\Phi_t, k_{it}, n_{it}, \omega_{it}) \tag{11}$$

The short-run profit together with firm's draw from the sunk cost and fixed cost distributions and the evolution of productivity determine firm's decisions to import, which I will explain in the next subsection.

2.3 Evolution of productivity and learning-by-importing effect

I specify the evolution of firm's productivity parametrically. Importing intermediate inputs from abroad would allow for a firm to have close contact with foreign suppliers in developed countries. This will lead to the positive dynamic externalities, which is referred to *learning-by-importing* effect. I assume that previous period import status, which is binary variable $d_{it-1} \in \{0, 1\}$, will affect today's productivity regardless of the intensive margin of import. Due to this learning-by-importing effect, firms may choose to import if they expect their productivity to grow quickly with importing even though it is not profitable in the static sense. I will explain this in detail in the next subsection.

I suppose that productivity ω_{it} evolves overtime as a first-order Markov process that depends on the previous productivity and the firm's import decision, which is approximated by the cubic polynomial:

$$\begin{aligned} \omega_{it} &= g(\omega_{it-1}, d_{it-1}) + \xi_{it} \\ &= \alpha_0 + \sum_{s=1}^{3} \alpha_s (\omega_{it-1})^s + \alpha_4 d_{it-1} + \xi_{it} \end{aligned} \tag{12}$$

where α_4 captures learning-by-importing effect. ξ_{it} is i.i.d. shock with mean 0 and variance σ_ξ^2 which is assumed to be independent of previous productivity and previous import.

2.4 Firm's dynamic decision on importing

First I will explain the timing assumption of the firm's decisions. At the beginning of each period, firm i observes the current productivity shock. Firm's state vector $\mathbf{s}_{it}$ is,

$$s_{it} = (\omega_{it}, d_{it-1}, k_{it}, \Phi_t, \mathbf{w}_t) \tag{13}$$

which includes today's productivity ω_{it}, previous period import status d_{it-1}, predetermined capital stock k_{it}, the market aggregate variables Φ_t, and factor prices $\mathbf{w}_t$. In this paper, I treat capital stock is fixed during the period of analysis as in Aw et al. (2011) and Bai et al. (2017).

After observing the state vector, a firm observes relevant cost of import and decides whether to import or not. If a firm is previous non-importer (i.e., $d_{it-1} = 0$), the firm observe startup sunk cost. If a firm is previous importer (i.e., $d_{it-1} = 1$), the firm observe fixed cost but not sunk cost. Introduction of two distinct costs depending on previous import status is motivated by the observed pattern of import participation which reveal substantial persistence. These costs will differ across firms because of differences in technological opportunities and expertise. Since the underlying technological opportunities are not observed, I model these costs as i.i.d. draws from a known distribution, which is the key objects in the second stage estimation. As I explain in detail below, making discrete import decision requires for a firm to compare the benefit from import and sunk or fixed cost of import. The benefit consists of static component (cost reduction effect) and dynamic component (learning-by-importing effect) and the former depends on the intensive margin of import (n_{it}). Therefore, in order to determine the discrete import choice, the intensive margin of import needs to be determined at this point. Since the model does not have an apparatus to pin down the intensive margin endogenously, as an interim solution, I impose additional assumption that the intensive margin of import is pre-determined by the firm size (capital stock) and there is no optimization problem that the firm solves in determining n_{it}.[2] Upon making discrete import decision, firms decides a total intermediate input expenditure, which is completely static, freely adjustable inputs.[3]

With this timing assumption in mind, I will provide a complete explanation on firm's discrete import decision. Firms draw relevant sunk or fixed cost of import depending on the previous period import status from a known joint distribution G^γ. Firms pay only the sunk cost (not the fixed cost) when switching to importer and only the fixed cost (not the sunk cost) when remaining as importer. Given firm's state vector s_{it}, value function before observing fixed cost and sunk cost is, by choosing $d_{it} = \{0, 1\}$,

$$V(s_{it}) = \int \max_{d_{it}} \Big\{ u(d_{it}, s_{it}|\gamma_{it}) + \delta \mathbb{E}_t \left[V(s_{it+1}, d_{it}) \right] \Big\} dG^\gamma \tag{14}$$

where $u(d_{it}, s_{it}|\gamma_{it})$ is current period payoff that depends on the choice of import, the current state, and the realization of the relevant import expenditure. As formally defined below, current period

[2]More detailed explanation is provided in section 3.2.

[3]In my model specification, importing may reduce the cost of input bundle. Therefore, we can view that importing firms may purchase the intermediate input with lower unit cost.

payoff depends on marginal cost which depends on state variables and the intensive margin of import n_{it}. However, the intensive margin of import is assumed to be determined by the capital stock. Therefore, $u(\cdot)$ does not include n_{it} as its argument. For the notational simplicity, let $d_{it}^N = (1 - d_{it})$ and $d_{it}^I = d_{it}$. Current period payoff conditional on relevant import cost γ_{it} can be written as:

$$u(d_{it}, s_{it}|\gamma_{it}) = d_{it}^N \Pi_{it}^N + d_{it}^I \left[\Pi_{it}^I - (d_{it-1}^N \gamma_{it}^S + d_{it-1}^I \gamma_{it}^F)\right] \tag{15}$$

where operational profit before paying the fixed or sunk cost is import-decision-specific, namely:

$$\Pi_{it}^m = \frac{1}{\sigma} R_{it}^m = \left[\frac{\sigma}{\sigma - 1} MC_{it}\left(\omega_{it}, k_{it}, \mathbf{w}_{it} | d_{it}^m = 1\right)\right]^{1-\sigma} \frac{Y_t}{P_t^{1-\sigma}}, \quad m = \{N, I\} \tag{16}$$

The second component of firm's value function is expected continuation value $\mathbb{E}_t\left[V(s_{it+1}, d_{it})\right]$ where expectation is taken over next period productivity:

$$\mathbb{E}_t\left[V(s_{it+1}, d_{it})\right] = \int_{\omega'} V(s') dF(\omega' | \omega_{it}, d_{it}) \tag{17}$$

where $F(\omega'|\omega_{it}, d_{it})$ is defined by the productivity evolution as in equation (12). For any state vector s_{it+1}, denote the choice specific continuation value from choosing $d_{it}^m = 1$ for $m = \{N, I\}$ as $\mathbb{E}_t V_{it+1}^m = \mathbb{E}_t\left[V(s_{it+1}, d_{it}^m = 1)\right]$.

Firm's import decisions depend on the difference in the pairwise marginal benefits between two options (importing and not importing) and the associated fixed or sunk cost of import. Let Δ_{it} be marginal benefit of being an importer versus being a non-importer, which has two components:

$$\Delta_{it} = \Pi_{it}^I - \Pi_{it}^N + \delta\left[\mathbb{E}_t V_{t+1}^I - \mathbb{E}_t V_{t+1}^N\right] \tag{18}$$

This implies that even though a firm finds it not profitable to import in a static sense, importing yields positive gain if it increases expected continuation value through the learning-by-importing effect and potentially lower fixed cost to be incurred in the future.

Given the distributions of sunk and fixed cost of import, the marginal benefit and import status in last period pin down the switching probability. For any set of realized import cost $\boldsymbol{\gamma}_{it} = \left(\gamma_{it}^{S}, \gamma_{it}^{F}\right)$, I can calculate the differences in life-time payoff between non-importer and importer, which is unconditional on the previous period import status.

$$y_{it}^{IN} = \Delta_{it} - \mathbb{1}\{d_{it-1}^{N} = 1\}\gamma_{it}^{S} - \mathbb{1}\{d_{it-1}^{I} = 1\}\gamma_{it}^{F} \tag{19}$$

where $\mathbb{1}\{\cdot\}$ is an indicator function. Therefore, the unconditional probability to become importer is given by:

$$\Pr\left(y_{it}^{IN} \geq 0\right) = \Pr\left(\mathbb{1}\{d_{it-1} = N\}\gamma_{it}^{S} + \mathbb{1}\{d_{it-1} = I\}\gamma_{it}^{F} \leq \Delta_{it}\right) \tag{20}$$

Conditioning on the previous import status, the probability of changing (or not changing) the import status can be expressed as follows. For example, P_{it}^{NI} is the probability that firm i which is non-importer (N) in period $t-1$ becomes importer (I) in period t. For a previous non-importer, the probability to stay as non-importer and to switch to importer, respectively, are given by:

$$\begin{aligned} P_{it}^{NN} &= \Pr\left(y_{it}^{IN}|_{d_{it-1}^{N}=1} \leq 0\right) = \Pr\left(\gamma_{it}^{S} \geq \Delta_{it}\right) \\ P_{it}^{NI} &= \Pr\left(y_{it}^{IN}|_{d_{it-1}^{I}=1} > 0\right) = \Pr\left(\gamma_{it}^{S} < \Delta_{it}\right) \end{aligned} \tag{21}$$

Therefore, we can write the Bellman's equation for the previous non-importer as follows:

$$V_{it}^{N} = P_{it}^{NN}\left(\Pi_{it}^{N} + \delta\mathbb{E}_{t}V_{it+1}^{N}\right) + P_{it}^{NI}\left(\Pi_{it}^{I} - \mathbb{E}\left[\gamma_{it}^{S}|\gamma_{it}^{S} < \Delta_{it}\right] + \delta\mathbb{E}_{t}V_{it+1}^{I}\right) \tag{22}$$

Note that superscript N of V_{it}^{N} indicates the previous period import status. The expectation of sunk cost is conditional on $\gamma_{it}^{S} < \Delta_{it}$ since this is the only case which is relevant for the switching firm. As in Aw et al. (2011) and Bai et al. (2017), I assume that sunk and fixed cost is drawn from exponential distribution with parameter λ^{m} for $m \in \{S, F\}$. With this distributional assumption, I can simplify the value function as follows,

$$V_{it}^{N} = \Pi_{it}^{I} + \delta \mathbb{E}_t V_{it+1}^{I} - \lambda^S G^S(\Delta_{it})$$

Appendix A outlines the step-by-step derivation of the value function.

Analogously, we can get the probability to stay as importer and to switch to non-importer for a previous importer:

$$\begin{aligned} P_{it}^{II} &= \Pr\left(y_{it}^{IN}|_{d_{it-1}^{I}=1} > 0\right) = \Pr\left(\gamma_{it}^{F} < \Delta_{it}\right) \\ P_{it}^{IN} &= \Pr\left(y_{it}^{IN}|_{d_{it-1}^{I}=1} \leq 0\right) = \Pr\left(\gamma_{it}^{F} \geq \Delta_{it}\right) \end{aligned} \tag{23}$$

The Bellman's equation for a previous importer can be written as

$$V_{it}^{I} = P_{it}^{II}\left(\Pi_{it}^{I} - \mathbb{E}\left[\gamma_{it}^{F}|\gamma_{it}^{F} < \Delta_{it}\right] + \delta \mathbb{E}_t V_{it+1}^{I}\right) + P_{it}^{IN}\left(\Pi_{it}^{N} + \delta \mathbb{E}_t V_{it+1}^{N}\right) \tag{24}$$

Finally, I assume that firms make draw from the sunk and fixed costs distribution each period independently.

With this setup, we would expect that, other things equal, if a fixed cost of import is higher, then continuation value of being importer is lower, which decreases the marginal benefit of importing. As for sunk cost of import, if it is higher, continuation value of being non-importer is lower, which increases the marginal benefit of importing. Finally, if a learning-by-importing effect is better, then continuation value of being importer is greater, which increases the marginal benefit of importing. Therefore, these three components governs firm's discrete decision on import in this model.

3 Empirical Strategy

This section outlines the empirical strategy to estimate the structural model of firm's import decision developed in the previous section.

3.1 Stage 1: Elasticity and productivity

Each firm's total variable cost can be expressed as

$$
\begin{aligned}
TVC_{it} &= MC_{it}q_{it} \\
&= \frac{\sigma-1}{\sigma}P_{it}Q_{it} \\
&= \frac{\sigma-1}{\sigma}R_{it}
\end{aligned} \tag{25}
$$

Since the total variable cost and revenue are data, I can estimate the following equation by the ordinary least squares (OLS),

$$
TVC_{it} = \zeta R_{it} + \nu_{it} \tag{26}
$$

where ν_{it} is a measurement error. I compute the total variable cost as the sum of wage payable and intermediate input purchase for production. R_{it} is total sales income (revenue). Elasticity of substitution is recovered by $\hat{\sigma} = 1/(1-\hat{\zeta})$

Next I estimate the log revenue function defined by (10). Let $\phi_0 = (1-\sigma)\ln\left(\frac{\sigma}{\sigma-1}\right) + (1-\sigma)\gamma_0$ and $\phi_t = \ln\Phi_t + (1-\sigma)\beta_t$. Econometric specification for the log revenue function is

$$
\ln(r_{it}) = \phi_0 + \sum_{t=1}^{T}\phi_t D_t + (1-\sigma)(\beta_k k_{it} + \beta_n n_{it} - \omega_{it}) + v_{it} \tag{27}
$$

where v_{it} is a measurement error. Following the insights of Olley and Pakes (1996), I rewrite the unobserved productivity in terms of observables that are correlated with productivity. In particular, I employ Levinsohn and Petrin's (2003) method and control the unobserved productivity using the fact that more productive firms will use more materials. Since marginal cost is constant in output, the relative expenditures on all the variable inputs will not be a function of total output. The assumption that productivity is Hicks neutral is also important, since, otherwise, the difference in productivity leads to variation across firms and time in the mix of variable inputs used. I assume that productivity is a function of capital stock and material inputs, $\omega^*_{it}(k_{it}, x_{it})$. Plug this into equation (27) and get

$$
\begin{aligned}
\ln(r_{it}) =& \phi_0 + \sum_{t=1}^{T} \phi_t D_t + (1-\sigma)\left(\beta_k k_{it} + \beta_n n_{it} - \omega^*_{it}(k_{it}, x_{it})\right) + v_{it} \\
=& \phi_0 + \sum_{t=1}^{T} \phi_t D_t + h(k_{it}, x_{it}, n_{it}) + v_{it}
\end{aligned} \tag{28}
$$

where $h(\cdot)$ captures the combined effect of capital, intensive margin of import, and productivity on revenue. I approximate $h(\cdot)$ by a 3rd-degree polynomial of its argument and estimate equation (28) by the OLS.

I denote by $\widehat{\phi}_{it}$ a fitted value of $h(\cdot)$, which is an estimator of $(1-\sigma)(\beta_k k_{it} + \beta_n n_{it} - \omega^*(k_{it}, x_{it}))$. I then construct a productivity series for each firm. This is done by estimating the parameters of productivity process specified in equation (12). Substituting $\omega_{it} = \frac{-1}{1-\sigma}\widehat{\phi}_{it} + \beta_k k_{it} + \beta_n n_{it}$ into the productivity evolution equation gives an estimating equation,

$$
\begin{aligned}
\widehat{\phi}_{it} =& (1-\sigma)\left(\beta_k k_{it} + \beta_n n_{it}\right) - (1-\sigma)\left[\alpha_0 + \sum_{s=1}^{3} \alpha_s (\omega_{it-1})^s + \alpha_4 d_{it-1} + \xi_{it}\right] \\
=& \beta^*_k k_{it} + \beta^*_n n_{it} - \alpha^*_0 + \alpha_1\left(\widehat{\phi}_{it-1} - \beta^*_k k_{it-1} - \beta^*_n n_{it-1}\right) - \alpha^*_2\left(\widehat{\phi}_{it-1} - \beta^*_k k_{it-1} - \beta^*_n n_{it-1}\right)^2 \\
&+ \alpha^*_3\left(\widehat{\phi}_{it-1} - \beta^*_k k_{it-1} - \beta^*_n n_{it-1}\right)^3 - \alpha^*_4 d_{it-1} - \xi^*_{it}
\end{aligned} \tag{29}
$$

where the star represents that the coefficients are multiplied by $(1-\sigma)$, with two exceptions, $\alpha^*_2 = (1-\sigma)^{-1}\alpha_2$ and $\alpha^*_3 = (1-\sigma)^{-2}\alpha_3$. Equation (29) can be estimated by the nonlinear least square. The variance of error term, σ^2_ξ, is pinned down by the sample variance of the residuals.

It is worth noting that my timing assumption let a firm to determine the intensive margin of import n_{it} and intermediate purchase x_{it} after observing the current productivity shock. This would generate the potential endogeneity problem between these two variables and current productivity shock ξ_{it}. In my current manuscript, I do not control for the endogeneity of intermediate input. As for intensive margin of import, I use mean level of n_{it} conditional on firm size (capital stock), instead of using the actual value observed in the data. First, I classify all firm-year observations into eight bins based on the capital stock. These bins are created such that each bin contains the same number of observations. More specifically, I compute the every 12.5 percentile value of capital stock in the sample, denoted by $\tilde{k}_j$ for $j = 0, ..., 8$ with $\tilde{k}_0 = \min\{k_{it}\}$ and $\tilde{k}_8 = \max\{k_{it}\}$, and define the each bin of capital stock as $\kappa_j = [\tilde{k}_{j-1}, \tilde{k}_j)$ for $j = 1, ..., 8$. [4] For each bin j, I compute the average of n_{it} as follows,

$$\overline{n}_j = \frac{1}{\#\{it|k_{it} \in \kappa_j, d_{it} = 1\}} \sum_{it \in \{it|k_{it} \in \kappa_j, d_{it}=1\}} n_{it} \tag{30}$$

I use this conditional mean of n_{it} in estimating equation (29). Using the conditional mean in place of actual value of n_{it} would partially aid the potential endogeneity problem mentioned above. Furthermore, it also helps in solving the firm's value function in the second stage estimation, which I will explain below.

3.2 Stage 2: Dynamic estimation

I follow Aw et al. (2011) for the basic framework to estimate the dynamic parameters in the second stage. In this stage, I need to solve firm's value function where state variables include continuous variables, i.e., capital and productivity. I discretize the state space as follows: in discretizing capital stock variable, I generate eight bins of capital stock, κ_j for $j = 1, ..., 8$, in the way I explained in the previous subsection.[5] For each bin, I compute the mean, $\overline{k}_j = \frac{1}{\#\{it|k_{it} \in \kappa_j\}} \sum_{\{it|k_{it} \in \kappa_j\}} k_{it}$, which gives eight grid points of capital stock $\overline{k}_j$ for $j = 1, .., 8$. Since the model does not endogenize firm's investment decision, in the second stage of estimation, I assume that capital stock is fixed over time. In assigning the relevant grid point of capital stock to each firms in the sample, I compute the mean capital stock for each firm over the sampled period and assign the relevant grid point.

As for discretizing productivity, I follow Rust's (1987) random grid approach. First, I draw 100 random numbers from the halton sequence, denoted by u_h for $h = 1, 2, ..., 100$. Let $\omega^{max} \equiv \max\{\omega_{it}\}$ and $\omega^{min} \equiv \min\{\omega_{it}\}$, respectively, denote the maximum and minimum values of productivity which is constructed in the first stage. Random grid point for productivity is constructed by $\omega_h = \omega^{min} + (\omega^{max} - \omega^{min}) \times u_h$ for $h = 1, 2, ..., 100$.

Parameters to be estimated in the second stage are exponential distribution parameters for fixed and sunk cost of import. I estimate the parameters by directly maximizing the likelihood and by employing the Bayesian MCMC. In both methods, I need to compute the likelihood of the data given parameters. In order to obtain the likelihood of the data, I need to solve firm's value function

[4] These bins are also used in discretizing the state space for capital stock in the second stage estimation.

[5] The baseline result is based on the discretization of state space for capital with 8 grid points. I also implemented the second stage estimation with finer grids (16 grid points).

to compute the conditional choice probability. I employ the nested fixed point algorithm *à la* Rust (1987).

(1) Begin with an initial guess of parameters $G^{\Upsilon^0} = (\lambda^{S0}, \lambda^{F0})$ where $\lambda^m > 0$, $m = \{S, F\}$, is the exponential distribution parameter for sunk/fixed cost of import.

(2) Calculate the transition probability of productivity $F(\omega'|\omega, d)$ based on estimated parameters in the first stage. For example, suppose that today's productivity is ω and import status is d. From equation (12), I know that $\widehat{\xi}(\omega, d, \omega') \equiv \omega' - \left(\widehat{\alpha}_0 + \sum_{s=1}^{3} \widehat{\alpha}_s(\omega_h)^s + \widehat{\alpha}_4 d\right)$ is distributed normal distribution with mean 0 and variance $\widehat{\sigma}_\xi^2$ where the hat-notation indicates that they are obtained in the first stage estimation. Let $\varphi(\cdot|0, \widehat{\sigma}_\xi^2)$ be the density function of normal random variable with mean 0 and variance $\widehat{\sigma}_\xi^2$. Consider the transition probability from ω to ω'_h, where ω'_h is the hth random grid point. I can compute the transition probability as:

$$\Pr(\omega'_h|\omega, d) = \frac{\varphi(\widehat{\xi}(\omega, d, \omega'_h)|0, \widehat{\sigma}_\xi^2)}{\sum_k \varphi(\widehat{\xi}(\omega, d, \omega'_k)|0, \widehat{\sigma}_\xi^2)}$$

This transition probability is used to calculate the expected continuation value.

(3) Iterate on the following inner loop to find fixed points of value functions

(a) Begin with a set of initial guess of value functions $V^{m0}(s)$ for $m = \{N, I\}$. State variables are triplet, capital, productivity, and previous import status. With discretized state space, I have two sets of 100×8 matrices of value functions, each set corresponds to the value functions of previous non-importer and previous importer.

(b) Using the transition probability obtained in the previous stage, calculate the expected continuation value as non-importer and importer, where expectation is taken over the next period productivity:

$$\mathbb{E}V^{m0} = \int_{\omega'} V^{m0}(s')dF(\omega'|\omega, d^m) = \sum_h V^{m0}(k', d^m, \omega'_h) \Pr(\omega'_h|\omega, d^m)$$

(c) Update value functions for importer and non-importer according to equation (22) and (24). Choice probability on import is defined in equation (21) and (23). For example, in the case of non-importer's value function:

$$V^{N1} = P^{NN}\left(\Pi^N + \delta\mathbb{E}V^{N0}\right) + P^{NI}\left(\Pi^I - \mathbb{E}\left[\gamma^S|\gamma^S < \Delta IN\right] + \delta\mathbb{E}V^{I0}\right)$$

Calculating the importer's profit requires the intensive margin of import n. In the current manuscript, the model does not endogenize the firm's decision on intensive margin of import. Therefore, I assume that a firm with capital stock $\overline{k}_j$ imports by the amount $\overline{n}_j$ for $j = 1, ..., 8$ which is defined in equation (30). [6]

(d) Iterate the step (a) – (c) until $|V^{m,n+1} - V^m| < \varepsilon$ where ε is predetermined tolerance level.

(4) Given the value function, evaluate the likelihood function of every observation. For example, let's consider a observation with $(\overline{k}_j, \omega, d^N)$, i.e., a firm which has capital stock $\overline{k}_j$ (on grid point), productivity ω, and previous non-importer status. Using the value function solved in the previous step, I can compute this firm's choice probability to remain as non-importer and to become importer in the next period, denoted by P^{NN} and P^{NI}, respectively. Let S^{NN} (S^{NI}, respectively) be the dummy variable which takes 1 if a firm becomes non-importer (importer) this period conditional on being non-importer last period. Contribution of this observation to the likelihood is $(S^{NN})^{P^{NN}} + (1 - S^{NN})^{P^{NI}}$.

In addition to the maximum likelihood estimation, I implement the Bayesian Markov Chain Monte Carlo estimation as well. The detailed procedure is explained in Appendix C.1

4 Empirical Results

4.1 Data

In the empirical analysis, I use two kinds of firm-level data in China. The first one consist of firm-level data from the Annual Survey of Industrial Production (ASIP) conducted by the Chinese government's National Bureau of Statistics. This survey includes all state-owned enterprises (SOEs)

[6] The underlying assumption is that intensive margin of import is (monotonically) increasing in capital stock, i.e., the larger a firm is, the more the firm uses imported varieties. However, computed values of $\overline{n}_j$ do not reveal the expected relationship to the size of capital: $\overline{n}_1 = 0.0439$, $\overline{n}_2 = 0.0471$, $\overline{n}_3 = 0.0203$, $\overline{n}_4 = 0.0425$, $\overline{n}_5 = 0.0521$, $\overline{n}_6 = 0.0351$, $\overline{n}_7 = 0.0301$, and $\overline{n}_8 = 0.0265$. In the literature, Gopinath and Neiman (2014) and Halpern et al. (2015) argue that the importing firms may incur the fixed cost to increase the number of imported varieties. If this is the case, the capital stock as well as idiosyncratic productivity of the firm jointly determines n_{it}. In Appendix D, I will outline the theoretical and empirical strategies to endogenize the number of imported varieties in my model framework.

and non SOEs with sales over 5 million Yuan. The data contains information on the firm's industry of production, ownership type, capital stocks, intermediate inputs, and revenues. The second data set is Chinese Customs transaction-level data. It records the universe of monthly transaction by Chinese firms that participated in international trade. This custom data includes the value of each transaction (in US dollars) at the 8-digit Harmonized System (HS8) product level. Since our theoretical framework does not incorporate product-level import decision, for each firm, I aggregate over all the HS8 imports to get the firm's total import. Since the custom data covers the period from 2000 to 2006, this is the sampled period I use in the estimation. I focus on a particular industry defined by 2-digit of International Standard Industrial Classification (ISIC) in the estimation. After data cleaning and taking into consideration the number of observations, I restrict our attention to privately owned firms in machinery industry (ISIC2 29). I use balanced panel sample, i.e., restricting the sample to the firms which appears in every year. For more detail about data cleaning, please refer to Appendix B

4.2 Summary statistics and empirical regularities

Table 1 shows the summary statistics of some key variables. Focusing on fraction of importer, among all firm-year observations in the sample, 11.3% is importers. Among 1,889 firms in the sample, 25.1% of them uses imported inputs at least once during the sampled period. The fraction of importers increases over the sampled period: in 2000 only 5.2% of firms uses imported inputs and the fraction increases to 15.1% in 2006. After estimating parameter, I will examine whether the simulated data implied by the parameter estimates generates the observed pattern of import participation rate.

Table 2 shows the transition of import status. Each row indicates current period status and each column indicates next period status. For example, if a firm is non-importer in the current period, 95.13% of them remains as non-importer in the next period. The table reveals the strong persistence of the import status displayed in diagonal elements. The table implies that a firm face a substantial sunk cost to enter the import market. While the current importers are more likely to remain as importer in the next period with probability 74.61%, more than 25% of them may stop importing next period. This implies fixed cost of import is not negligible although it would be smaller than sunk cost of import. The table confirms the necessity to include sunk cost and fixed cost separately in analyze firm's dynamic decision on import.

Table 1: Summary Statistics

	Log capital	Log Revenue	Fraction of Importer		
			Average	2000	2006
Num Obs	13,223	13,223	13,223	1,889	1,889
Mean	8.538	10.288	0.113	0.052	0.151
Median	8.477	10.130	0	0	0
Min	0.497	1.792	0	0	0
Max	14.933	16.780	1	1	1
Std Dev	1.526	1.201	0.434	0.222	0.358

Table 2: Transition Probability of Importing Status

$t \backslash t+1$	Nonimporter	Importer
Nonimporter	95.13%	4.87%
Importer	25.39%	74.61%

Note: Sample size 13,223.

4.3 Stage 1: elasticity, production function, and productivity evolution

In the first step, I begin with estimating the elasticity of substitution σ by running the ordinary least squares (OLS) for the regression equation (26). OLS estimate for ζ is presented in Table 3. This result implies that elasticity of substitution is 7.949.

Next, I estimate the log revenue function (27). $h(\cdot)$ is approximated by the third-degree polynomial in its arguments. Using the fitted value of $h(\cdot)$, estimate the equation (29) by the nonlinear least squares. Table 4 demonstrates the estimated coefficients. Since I use the lagged variables, number of observations is smaller than in the previous regression. The key parameters are β_n and α_4, which reflect the static cost reduction effect and learning-by-importing effect, respectively. β_n is negative and statistically significant. This implies that a importing firm may have lower marginal cost because it may use the more intermediate varieties compared with non-importer. α_4 is significantly positive, implying there is an dynamic productivity implication from importing. Therefore, the first stage estimation confirms the two benefits from using imported intermediate inputs.

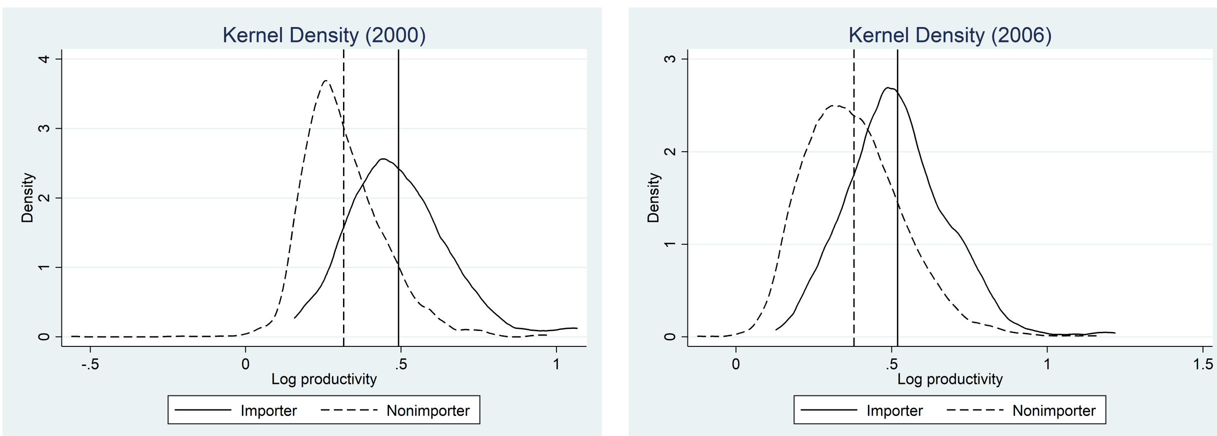
Given the estimates, I can recover the series of productivity ω_{it}. Table 5 demonstrates the summary statistics of estimated productivity. On average, productivity of importer is higher than non-importer. Mean productivity is .367 for non-importer and .507 for importer. This is consistent with the empirical regularities found in the previous studies, such as Bernard et al. (2012). Figure 1 shows kernel density estimates of productivity for non-importers and importers in 2000 and 2006. Solid and dashed vertical lines, respectively, show the mean productivity of importer and

Table 3: Elasticity of Substitution

	(1)
	Total Variable Cost
Total Sales Income	0.874***
	(0.001)
Observations	13,223
R^2	0.983

Standard error in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 1: Productivity distribution



non-importer. The figure reveals that productivity distributions for importers and non-importers are substantially overlapping. It implies that productivity sorting of importers and non-importers is not perfect in the data. This motivates to introduce the heterogeneous sunk and fixed cost of import.

4.4 Stage 2: fixed and sunk cost distributions

Baseline result

Table 6 summarizes the result for dynamic estimation. I implement the maximum likelihood and the Bayesian MCMC estimation. In Appendix C.2, I present the chains obtained by the MCMC algorithm and the posterior distributions of the parameters.

The table shows that the two methods give the different estimates. In order to assess the fit of the model to the data, I simulate the firm's import participation patterns over time given the parameter estimates. I simulate the import participation patterns of 1,889 firms, which is same

Table 4: Production function and productivity evolution

	Equation (29)
β_k	-0.016***
	(0.001)
β_n	-0.647*
	(0.066)
α_0	0.101***
	(0.005)
α_1	0.451***
	(0.028)
α_2	0.962***
	(0.055)
α_3	-0.531***
	(0.037)
α_4	0.012***
	(0.002)
Observations	11334
R^2	0.869

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

as the number of firms in the sample. Capital stock, first year productivity, and first year import participation status are taken from the data. Firstly, I solve the set of value functions given the parameter estimates. Then, I simulate the productivity of period $t \geq 2$ of each firm based on period $t-1$ productivity and import status according to the equation (12). For each firm, I draw the cost of import from the relevant exponential distribution depending of the previous period import status. By comparing the benefit and cost of importing, I can get the firm's binary decision of import at period t. Repeat this for 30 periods onward. Then, I compute the import participation rate by calculating the fraction of firms that import in each point in time.[7]

Figure 2 demonstrates the evolution of import participation rate among all firms which are calculated from the actual data and simulated data. The figure reveals that the fit of the model based on ML estimators is much poorer than the model based on MCMC estimates. It would suggest that likelihood function has multiple local maxima and the Matlab optimizer finds a local maximum but not the global maximum.

[7] In order to sweep out the simulation noises, I repeat the simulation 100 times and take average of the import participation rate.

Table 5: Estimated Productivity

	All		2000		2006	
	Nonimporter	Importer	Nonimporter	Importer	Nonimporter	Importer
Num Obs	11,733	1,490	1,791	98	1,604	285
Mean	.367	.507	.315	.492	.379	.519
Median	.326	.495	.293	.483	.365	.503
Min	-.561	.109	-.561	.157	-.124	.129
Max	1.169	1.217	.981	1.066	1.169	1.217
Std Dev	.141	.160	.131	.166	.158	.163

Table 6: Fixed/Sunk Cost Distribution Parameters (Baseline)

	MLE		Bayesian MCMC	
Parameter	Coefficient	Std Err	Mean	Std Dev
Fixed cost (λ^F)	2.0702	0.0288	3.1555	0.0627
Sunk cost (λ^S)	38.8186	2.0335	55.1810	2.6310

Finer grids: discretizing the state space for capital with 16 grid points

In the baseline estimation, I discretize the state space for capital with 8 grid points following the previous works. Given the variation in capital stock across firms, this might be too coarse to capture the heterogeneity in import decision among firms with different size of capital. Therefore, I re-estimate the model by doubling the number of grid points. Table 7 shows the estimation results. The Markov chains and posterior distributions for the two parameters are presented in Appendix C.3 . With finer grid points, the two alternative methods give almost the same estimates. Figure 3 demonstrate the trajectory of import participation based on the simulation. Comparing the result with Bayesian MCMC in the baseline model, the fit of the model improves slightly.

In my model framework, discretizing the state space with finer grids does not intensify the computational burden substantially. Time elapsed for ML estimation is 77.20, which is approximately 20 seconds longer than the baseline model. Therefore, implementing MLE with finer grid points would be more efficient than implementing the Bayesian MCMC with coarse grid in terms of computational time. But, of course, this is not the general result. The extent to which the computational time is affected by having finer grid points will depend on the model structure. Yet, this result would be suggestive in determining the grid size of the state space and the estimation method to be employed for the estimation of the other dynamic discrete choice problems. For example, contrary to the case here, if discretizing the state space with finer grids makes the computational time much longer, it would be better to have coarse grid points and implement MCMC.

Figure 2: Import Participation Rate (Baseline)

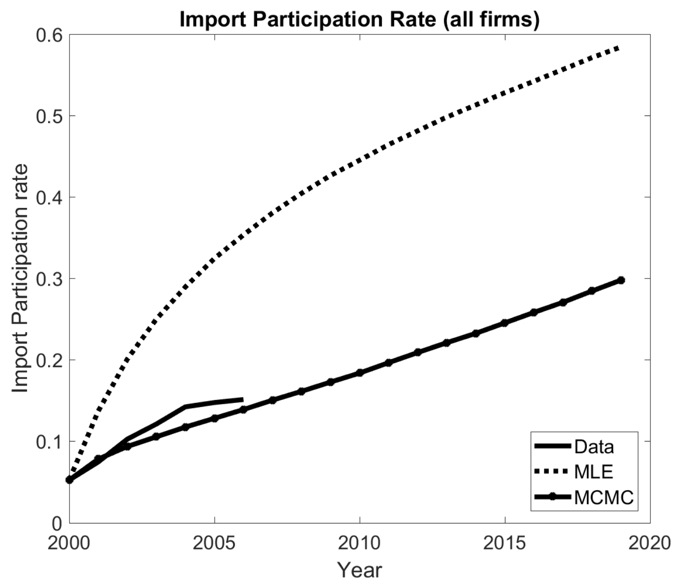


Table 7: Fixed/Sunk Cost Distribution Parameters (Finer Grids)

	MLE		Bayesian MCMC	
Parameter	Coefficient	Std Err	Mean	Std Dev
Fixed cost (λ^F)	3.3429	0.0488	3.3436	0.0488
Sunk cost (λ^S)	64.5385	3.3798	64.9167	3.1638

Different fixed and sunk costs depending on firm size

I also estimate the model which allows for the sunk and fixed cost distribution parameters to vary depending on the size of firms (i.e., capital stock). In this estimation, I use 16 grid points in discretizing the state space for capital stock. I categorize firms into 4 groups based on the size of capital stock. Size 1 is the group of firms with smallest capital stock and Size 4 is the group with largest capital stock. For example, firms which are assigned one of the lowest four grid points of the capital stock are labeled as Size 1 firms. I estimated the parameters through the Bayesian MCMC where initial guess on parameters are taken from the ML estimates. As explained in Appendix C.4, stationarity of the Markov chain for the sunk cost parameter of Size 1 firms is not very clear from the visual inspection of the generated chain. Therefore, it should be noted that the result presented below is relatively preliminary compared with the previous results.[8]

[8]As Figure C.3 in Appendix C.4 demonstrates, begin with the initial guess around 600, the chain starts to decline and it settles at the interval between 100 and 300 after 10,000 draws. It would be worth drawing more than 500,000 to determine whether the chain exhibits the stationarity.

Figure 3: Import Participation Rate (Finer Grids)

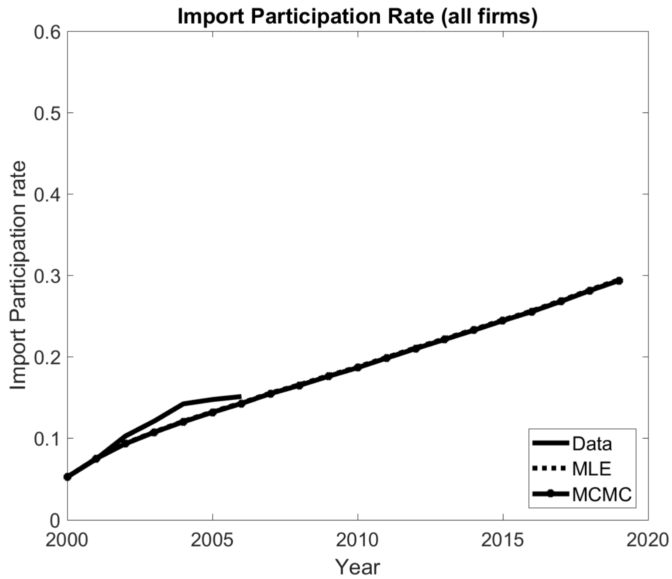


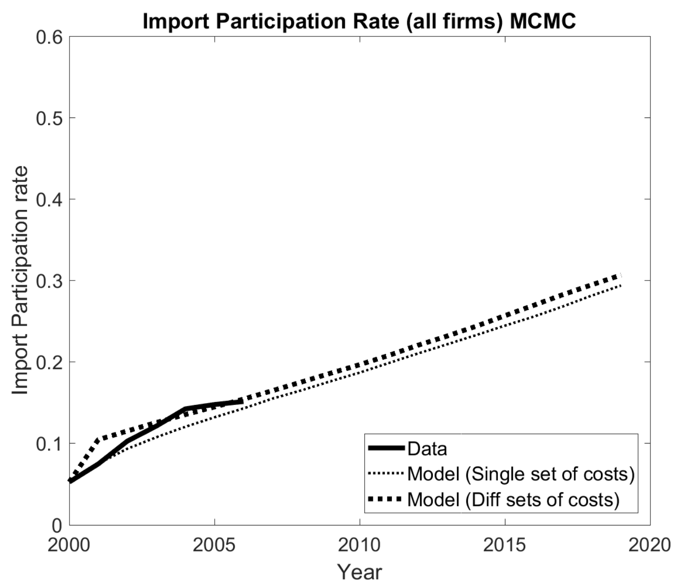
Table 8 shows the result. Firstly, estimated fixed cost parameters are smaller than the one obtained in the previous estimations. Second, while fixed cost of import does not reveal the substantial variation, sunk cost of import varies across firms with different size of capital stock. Typically, the sunk cost of import is higher for the small size firms and lower for larger firms.

Tabel D.1 demonstrates the fit of the estimated model to the data. For comparison, I presented the trajectory of import participation rate implied by the model estimated in section 4.4. Allowing for import expenditures to vary depending on firm size makes the fit of the model slightly better to the data. Figure C.4 in Appendix C.4 demonstrates the import participation rate for each category of firm size. Although the fit of the model improves for size 1 firms, it gets worse for size 4 firms. Therefore, this result would imply that there are another factors which govern the firm's import decisions. For example, in my model framework, sunk cost distribution is assumed to be fixed over time. However, along with the China's accession to the WTO in 2001, one may expect that there are some institutional reforms which affect the distribution of import cost itself. While it is beyond the scope of this paper, investigating the other determinants of firm's import decision would constitute an important question to be answered in the future study.

Table 8: Import Participation Rate (Finer Grids)

	Fixed Cost		Sunk Cost	
Size 1	1.4828	(0.2429)	155.7820	(28.3323)
Size 2	1.3880	(0.1551)	50.8662	(6.1751)
Size 3	1.5573	(0.1266)	37.0732	(3.5220)
Size 4	1.6648	(0.0796)	24.2712	(1.9162)

Figure 4: Import Participation Rate: Import Participation Rate: Heterogeneous Fixed/Sunk Cost



5 Counterfactual Analysis

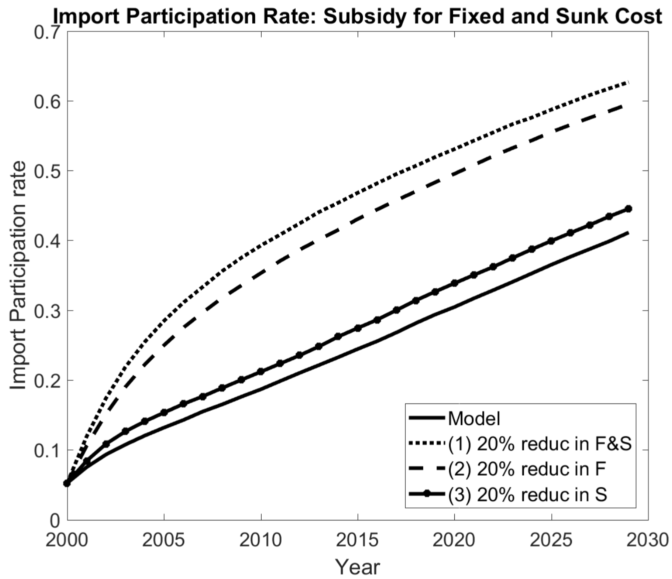
In this section, I will implement a simple counterfactual analysis to assess the effect of subsidizing firm's fixed or sunk cost of import. In the first stage estimation, I confirmed that the use of imported intermediate inputs contributes to increase the future productivity of firms via learning-by-importing effect. Hence, encouraging firms to start using imported inputs would be a potential policy apparatus to boost the aggregate productivity. Given this emphasis, in the counterfactual analysis, I considers the government subsidy scheme which is aimed at lowering the import expenditures. There are three schemes to be considered: (1) 20% reduction in fixed and sunk costs, (2) 20% reduction in fixed cost, and (3) 20% reduction in sunk cost.

This counterfactual analysis is based on the model discussed in section 4.4 which has the single set of fixed and sunk cost parameters and finer grid points in discretizing the state space of capital. Since the import cost is stochastically formulated in my model, I directly manipulate exponential distribution parameters. In the counterfactual world, fixed and sunk cost parameters are $\lambda^{F'} = 0.8\lambda^F$ and $\lambda^{S'} = 0.8\lambda^S$, respectively, which implies that import costs are lower by 20% *in expectation* than in the factual world. I assume that the subsidy is permanent.

Figure 5 shows the results. Solid line is the trajectory of import participation rate under the factual world. Dotted line labeled as (1) is the counterfactual with subsidy on both fixed and sunk cost of import. Dashed line labeled as (2) and solid line with dots labeled as (3) are, respectively, counterfactuals with subsidy only on fixed cost and on sunk cost. It is intuitive that subsidy on both fixed and sunk costs generate the highest import participation rate among the three schemes. On the other hand, it is somewhat surprising that subsidy on fixed cost is more effective in increasing the aggregate import participation rate than subsidy on sunk cost. After 30 years, import participation rate is 59.5% under subsidy on fixed cost and 44.5% under subsidy on sunk cost. Table 7 exhibits that the sunk cost of import is, on average, more than 19 times larger than the fixed cost of import. Therefore, one may expect that subsidizing sunk cost of import may plays more substantial role in enhancing import participation. However, as we see in Table 2, 25% of previous importers stops importing in the current period. Furthermore, firm's first-time entry decision to import involves not only sunk cost of import, but also fixed cost of import which is captured in the expected continuation value as an importer. Therefore, fixed cost of import also plays a significant role in determining firm's import decision both for previous non-importer and importer. And in the estimated model, the higher import participation rate is yielded under the subsidy scheme on fixed cost.

Although I do not discuss the cost effectiveness of these three schemes in the current manuscript, it is worth investigating which scheme would yield the highest import participation conditional on the government expenditure. Another interesting scope for the future study is the across-industry comparison. In this manuscript, I estimated the model using the samples of firms only in the machinery industry. I may estimate the model for other industries and compare the impact of import subsidy on the trajectory of import participation across sectors. These exercises would provide potentially important insights for policymakers in developing countries to design an effective policy scheme to encourage firms to use imported input, which in turn would contribute the increase in aggregate productivity of the economy.

Figure 5: Counter Factual Analysis: Import Cost Subsidy



6 Concluding Remarks

This paper develops and estimate the dynamic structural model of firm's import decision. The model features two benefits from using imported inputs; static cost reduction effect and dynamic learning-by-importing effect. The model allows heterogeneous startup sunk cost and fixed cost of import by introducing the stochastic formulations.

Using the Chinese firm-level data and custom data, I estimate the model with two-step procedure. In the first step, I estimat revenue production function and productivity evolution equation. In doing so, I control for the unobserved productivity by using the control function method *à la* Levinsohn and Petrin (2003). The estimation results confirm the evidence of the two benefits from using imported inputs. In the second stage, I estimat the dynamic parameters, i.e., fixed and sunk cost distribution parameters, by invoking the nested fixed point algorithm of Rust (1987). I estimated the parameters by directly maximizing the likelihood and by implementing the Bayesian MCMC estimation.

When I discretize the state space of capital with coarse grid, the MLE and the Bayesian MCMC estimation yield different results. By simulating the firm's import participation pattern implied by the parameter estimates, I assess the fit of the models estimated under the two alternative methods. The simulation reveals that the model based on MCMC estimates fits to the data much better than the MLE estimates. However, when I use the finer grid for the discretization of the state space, MLE and Bayesian MCMC give almost the same result. I also estimate the model which allows for the fixed and sunk cost of import to vary depending on firm size (capital stock). While the result is relatively preliminary, I confirmed the substantial variation in sunk cost across firms with different sizes. Sunk cost tends to be larger for smaller firms than larger firms. On the other hand, fixed cost of import seems not to depend on the firm size.

Finally, I implement the counterfactual analysis to assess the impact of import subsidy on the trajectory of aggregate import participation rate. I find that subsidy on fixed cost is more effective than subsidy on sunk cost to achieve the higher import participation rate.

References

Aw, Bee Yan, Mark J Roberts, and Daniel Yi Xu, "R&D Investment, Exporting, and Productivity Dynamics," *American Economic Review*, 2011, *101* (4), 1312–1344.

Bai, Xue, Kala Krishna, and Hong Ma, "How you export matters: Export mode, learning and productivity in China," *Journal of International Economics*, 2017, *104*, 122–137.

Bernard, Andrew B., J. Bradford Jensen, Stephen J. Redding, and Peter K. Schott, "The Empirics of Firm Heterogeneity and International Trade," *Annual Review of Economics*, 2012, *4* (1), 283–313.

Dixit, Avinash and Joseph E. Stiglitz, "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, 1977, *67* (3), 297–308.

Ethier, Wilfred J., "National and International Returns to Scale in the Modern Theory of International Trade," *American Economic Review*, 1982, *72* (3), 389–405.

Gopinath, Gita and Brent Neiman, "Trade Adjustment and Productivity in Large Crises," *American Economic Review*, 2014, *104* (3), 793–831.

Halpern, László, Miklós Koren, and Adam Szeidl, "Imported inputs and productivity," *American Economic Review*, 2015, *105* (12), 3660–3703.

Kasahara, Hiroyuki and Joel Rodrigue, "Does the use of imported intermediates increase productivity? Plant-level evidence," *Journal of Development Economics*, 2008, *87* (1), 106–118.

Levinsohn, James and Amil Petrin, "Estimating Production Funct Using Inputs to Control for Unobservables," *Review of Economic Studies*, 2003, *70* (70), 317–341.

Melitz, Marc J, "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 2003, *71* (6), 1695–1725.

Olley, G. Steven and Ariel Pakes, "The Dynamics of Productivity in the Telecommunictions Equipment Industry," *Econometrica*, 1996, *64* (6), 1263–1297.

Rust, John, "Optimal Replacement of GMC Bus Engines : An Empirical Model of Harold Zurcher," *Econometrica*, 1987, *55* (5), 999–1033.

A Deriving Value Function

This section outlines the derivation of equation (22) and (24). Here I will focus on non-importer's value function (derivation for importer's value function is analogous).

$$
\begin{aligned}
V_{it}^N &= P_{it}^{NN}\left(\Pi_{it}^N + \delta \mathbb{E}_t V_{it+1}^N\right) + P_{it}^{NI}\left(\Pi_{it}^I - \mathbb{E}\left[\gamma_{it}^S | \gamma_{it}^S < \Delta_{it}\right] + \delta \mathbb{E}_t V_{it+1}^I\right) \\
&= \Pr\left(\gamma^S > \Delta_{it}\right)\left(\Pi_{it}^N + \delta \mathbb{E}_t V_{it+1}^N\right) + \Pr\left(\gamma^S < \Delta_{it}\right)\left(\Pi_{it}^I - \mathbb{E}\left[\gamma_{it}^S | \gamma_{it}^S < \Delta_{it}\right] + \delta \mathbb{E}_t V_{it+1}^I\right) \\
&= \left(1 - \Pr\left(\gamma^S < \Delta_{it}\right)\right)\left(\Pi_{it}^N + \delta \mathbb{E}_t V_{it+1}^N\right) + \Pr\left(\gamma^S < \Delta_{it}\right)\left(\Pi_{it}^I - \mathbb{E}\left[\gamma_{it}^S | \gamma_{it}^S < \Delta_{it}\right] + \delta \mathbb{E}_t V_{it+1}^I\right) \\
&= G^S(\Delta_{it})\left[\Pi_{it}^I + \delta \mathbb{E}_t V_{it+1}^I - \Pi_{it}^N - \delta \mathbb{E}_t V_{it+1}^N\right] + G^S(\Delta_{it})\mathbb{E}\left[\gamma_{it}^S | \gamma_{it}^S < \Delta_{it}\right] + \Pi_{it}^N + \delta \mathbb{E}_t V_{it+1}^N \\
&= G^S(\Delta_{it})\Delta_{it} + G^S(\Delta_{it})\left[-\lambda^S + \frac{\Delta_{it}\exp(-\Delta_{it}/\lambda^S)}{G^S(\Delta_{it})}\right] + \Pi_{it}^N + \delta \mathbb{E}_t V_{it+1}^N \\
&= G^S(\Delta_{it})\Delta_{it} - \lambda^S G^S(\Delta_{it}) + \Delta_{it}\left[1 - G^S(\Delta_{it})\right] + \Pi_{it}^N + \delta \mathbb{E}_t V_{it+1}^N \\
&= \Pi_{it}^I + \delta \mathbb{E}_t V_{it+1}^I - \lambda^S G^S(\Delta_{it})
\end{aligned} \tag{A.1}
$$

where I compute the conditional expectation on γ^S as

$$
\mathbb{E}\left[\gamma^S | \gamma^S < \Delta_{it}\right] = \frac{\int \mathbb{1}\{\gamma^S < \Delta\}\gamma^S dG(\gamma)}{\int_0^{\Delta_{it}} dG^S(\gamma)} = \frac{\int_0^{\Delta_{it}} \gamma^S dG(\gamma)}{G^S(\gamma^S)} \tag{A.2}
$$

where

$$
\begin{aligned}
\int_0^{\Delta_{it}} \gamma^S dG(\gamma) &= \int_0^{\Delta_{it}} \frac{\gamma^S}{\lambda^S}\exp\left(-\frac{\gamma^S}{\lambda^S}\right) d\gamma^S \\
&= \lambda^S \int_0^{\Delta_{it}/\lambda^S} t\exp(-t)dt \\
&= \lambda^S \int_0^{\Delta_{it}/\lambda^S} t\left[-\exp(-t)\right]' dt \\
&= \lambda^S \left[\left[-t\exp(-t)\right]_0^{\Delta_{it}/\lambda^S} - \int_0^{\Delta_{it}/\lambda^S} (\exp(-t))dt\right] \\
&= \lambda^S\left[-\Delta_{it}/\lambda^S \exp(-\Delta_{it}/\lambda^S) - \exp(-\Delta_{it}/\lambda^S) + 1\right] \\
&= \lambda^S\left[1 - \exp(-\Delta_{it}/\lambda^S)\right] - \Delta_{it}\exp(-\Delta_{it}/\lambda^S) \\
&= \lambda^S G^S(\Delta_{it}) - \Delta_{it}\exp(-\Delta_{it}/\lambda^S)
\end{aligned} \tag{A.3}
$$

B Data description and cleaning

In the manufacturing census data, each observation must be identified by the unique pair of firm code (firm identification id) and year. However, among 1,559,993 observations, 119 pairs of ID and year have duplicate observations. I drop them from the sample. Census data records 4-digit ISIC (revision 3) code for each firm. I assign corresponding 2-digit ISIC code to each observation using the official concordance. There are 32 2-digit ISIC code in the data.

Custom data is monthly recorded and import values are in USD. I convert it to Chinese Yuan (RMB) using the exchange rate given in the dataset. I aggregate monthly data to obtain annual data.

Manufacturing census data and custom data is matched using the concordance for `firm_code` (firm ID in manufacturing census) and `party_id` (firm ID in custom data). Among 1,155,694 observations in the firm-level data, 216,300 observations are matched with the custom data. 405,015 observations in the custom data cannot be matched with firm-level data, which are dropped from the sample.

I will focus on the balanced panel sample. Among 505,507 firms in the sample, 10.2% of them have observations in every year. Therefore, the number of observations and firms, respectively, in balanced panel dataset is 361,102 and 51,586.

Among 51,586 firms in the data set, 2.34% of them switches the 2-digit ISIC code at least once during the sampled period. Among the ISIC switching firms, 89.58% of them switches once, 3.97% switches twice, and the rest switches more than twice. Since this paper focuses on the particular ISIC sector in the estimation, I make the ISIC code of a given firm fixed over the sampled period in the following manner: (i) if a firm has a unique ISIC code which appears most frequently during the sampled period (i.e., it has an unique mode), I replace with that code; (ii) if a firm has multiple modes, I replace with the ISIC code of the first year of the sampled year. Method (i) is applied to 98.92% the ISIC switching firms.

With respect to the firm ownership, the manufacturing census data classifies firms into 5 categories: (1) state owned enterprise; (2) collective enterprise; (3) foreign capital owned enterprise; (4) Hong Kong, Macao and Taiwan enterprise; and (5) privately owned enterprise. Among total 51,586 firms, 31.84% of them switch their ownership type at least once during the sampled period. Among 16,423 ownership type switching firms, 72.73% of them switches once, 18.58% switches twice, and the rest of them switches more than twice. In order to make the firm ownership type fixed over the sampled period, I again applied the same method as what is applied to the ISIC switching firms. Method (i) is applied to 98.62% of the ownership type switching firms.

I also drop the observations with negative sales income (0.63%), negative value of capital stock (0.88%), negative intermediate input purchase (0.68). 7.48% of observations reveal greater value of import than total intermediate input purchase. It implies that they purchase intermediate inputs abroad and resell them to other firms, which may be referred to *producer intermediary*. I drop these observations from the sample.

C Bayesian MCMC

C.1 MCMC Algorithm

I follow the Aw et al. (2011) for the implementation of the Bayesian MCMC estimation. Define the set of dynamic parameters as: $\Theta = (\lambda^F, \lambda^S, a_0, a_1, a_2)$, where a's are the parameters for logit equation for the initial conditions of importing. Using the model, compute the likelihood function defining the dataset (D) as a function of parameters $L(D|\Theta)$. Combining with a set of prior distributions $\pi(\Theta)$, the posterior distribution $P(\Theta|D)$ is well defined. Use MCMC techniques to calculate the moments of the posterior distribution. I adop diffuse priors for the parameters, namely, $N(0, 1000)$. The following outlines the Metropolis-Hastings algorithm I use:

(1) Draw $\mathcal{N} \times 2$ uniform random variables, where $\mathcal{N}$ is the number of repetition to construct the chain. Let μ_{np} be the representative element of the matrix of the random variables.

(2) Begin with the initial guess on the parameters: $\Theta^0 = (\boldsymbol{\lambda}^0, \mathbf{a}^0)$. I use the parameter estimates with MLE as the initial guess.

(3) Compute the prior density $\pi(\Theta^0)$ and likelihood $L(D|\Theta^0)$, which gives the posterior density by $\ln P(\Theta^0|D) = \ln(\pi(\Theta^0)) + \ln(L(D|\Theta^0))$

(4) Update parameter $\mathbf{a}^{prop} = \mathbf{a}^0 + s^a \times u$ where s^a is predetermined step size and u is an uniform random variable $u \sim Unif[0, 1]$

(5) Compute the posterior using the proposal: $\ln P(\boldsymbol{\lambda}^0, \mathbf{a}^{prop}) = \ln(\pi(\boldsymbol{\lambda}^0, \mathbf{a}^{prop})) + \ln(L(D|\boldsymbol{\lambda}^0, \mathbf{a}^{prop}))$

(6) Accept the proposal if $\ln(\mu_{1,1}) < \min\left\{0, \ln P(\boldsymbol{\lambda}^0, \mathbf{a}^{prop}) - \ln P(\Theta^0)\right\}$ and let $\mathbf{a}^1 = \mathbf{a}^{prop}$. Otherwise, $\mathbf{a}^1 = \mathbf{a}^0$

(7) Update parameter $\boldsymbol{\lambda}^{prop} = \mathbf{a}^0 + s^\lambda \times u$ where s^a is predetermined step size and u is an uniform random variable $u \sim Unif[0, 1]$

(8) Compute the posterior using the proposal: $\ln P(\boldsymbol{\lambda}^{prop}, \mathbf{a}^1) = \ln(\pi(\boldsymbol{\lambda}^{prop}, \mathbf{a}^1)) + \ln(L(D|\boldsymbol{\lambda}^{prop}, \mathbf{a}^1))$

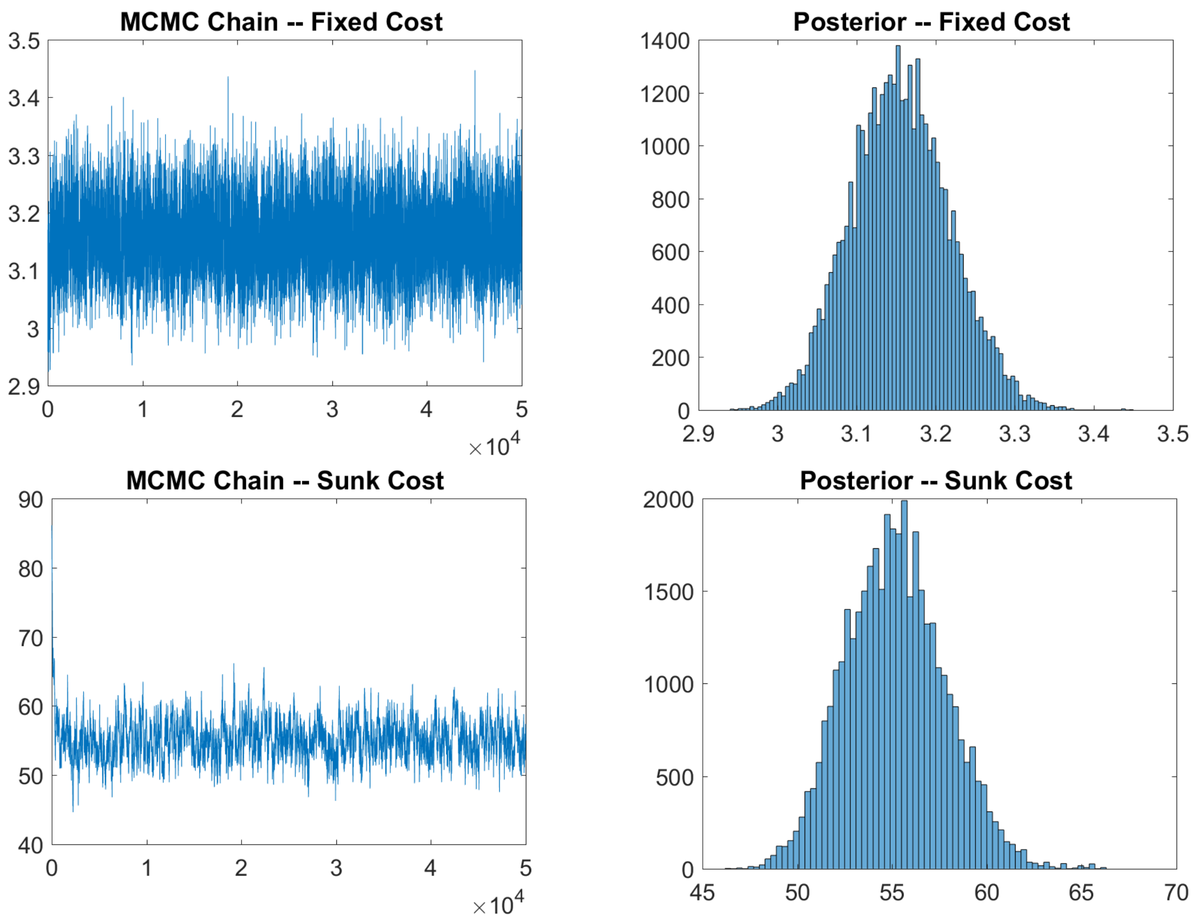
(9) Accept the proposal if $\ln(\mu_{1,2}) < \min\left\{0, \ln P(\boldsymbol{\lambda}^{prop}, \mathbf{a}^1) - \ln P(\boldsymbol{\lambda}^0, \mathbf{a}^1)\right\}$ and let $\boldsymbol{\lambda}^1 = \boldsymbol{\lambda}^{prop}$. Otherwise, $\boldsymbol{\lambda}^1 = \boldsymbol{\lambda}^0$

(10) Repeat (2)–(9) for $\mathcal{N}$ times.

C.2 Baseline Model

I set the number of repetition to $\mathcal{N} = 50,000$. In computing the posterior moments, I burn-in the first 10,000 draws. Figure C.1 demonstrates the chains obtained by the Metropolis-Hastings algorithm and the posterior density (after burning-in).

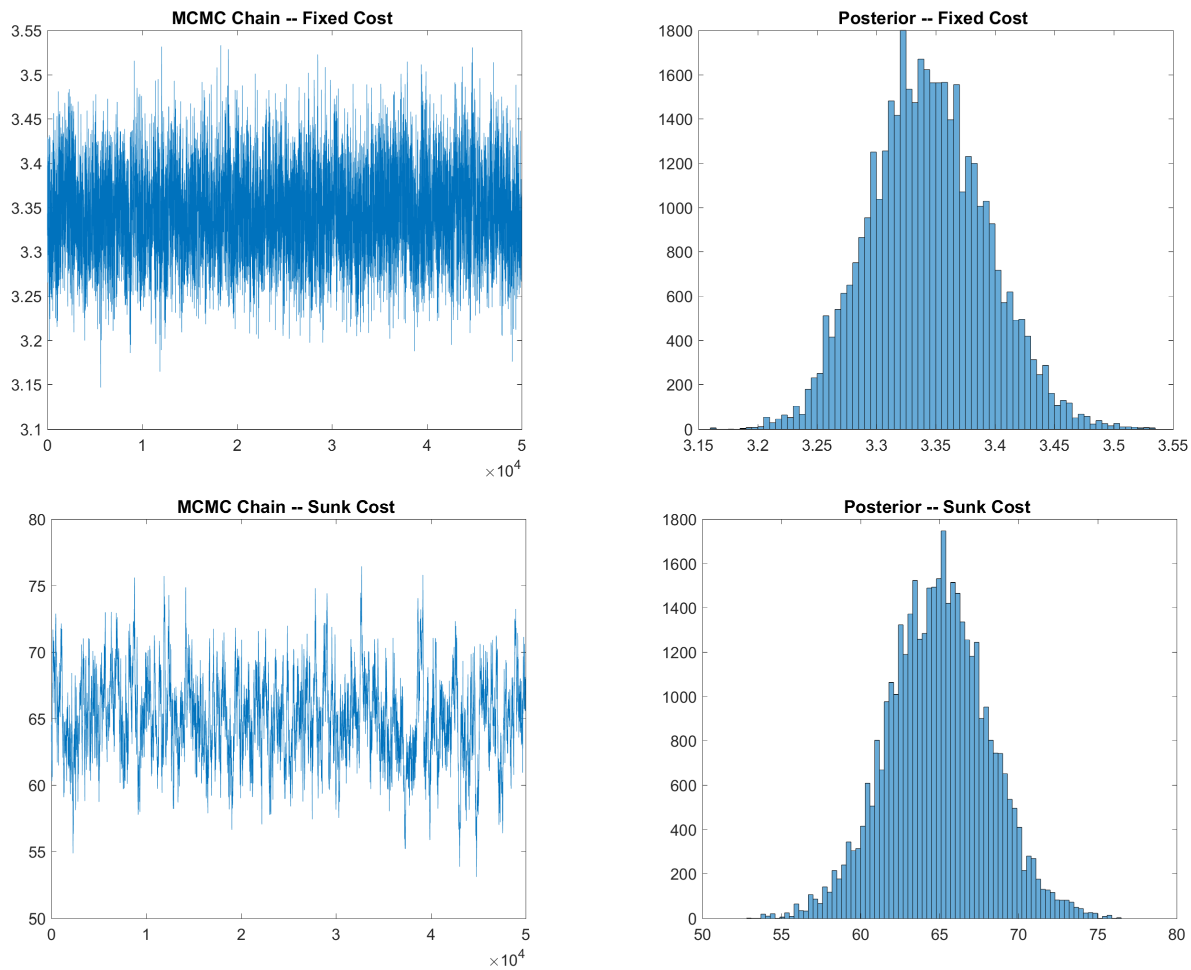
Figure C.1: Markov Chains and Posterior Density (Baseline)



C.3 Finer Grids (16 Grids Points)

In this estimation, state space for capital is discretized with 16 grid points. In implementing the MCMC, I set the number of repetition to $\mathcal{N} = 100,000$. In order to compute the posterior moments, I burn-in the first 10,000 draws. Figure C.2 demonstrates the chains obtained by the Metropolis-Hastings algorithm and the posterior density (after burning-in).

Figure C.2: Markov Chain and Posterior Density (Finer Grids)



C.4 Different Fixed/Sunk Cost Depending on Firm Size

State space for capital is discretized with 16 grid points. In implementing the MCMC, I set the number of repetition to $\mathcal{N} = 500,000$ and burn-in the first 25,000 draws in calculating the moments of posterior distributions. Figure C.2 demonstrates the chain obtained by the Metropolis-Hastings algorithm and the posterior density (after burning-in) for the sunk cost parameter for size 1 firm.[9] For this parameter, the stationarity of the chain is not very clear from the visual inspection. Therefore, the result for this estimation is relatively more preliminary to the other results.

Figure C.4 demonstrates the import participation rate implied by the model for each firm size category.

[9] In order to save space, the chains for other parameters are not presented. Visual inspection confirms the stationarity of the chains.

Figure C.3: Markov Chains and Posterior Density for Sunk Cost Parameter (Size 1)

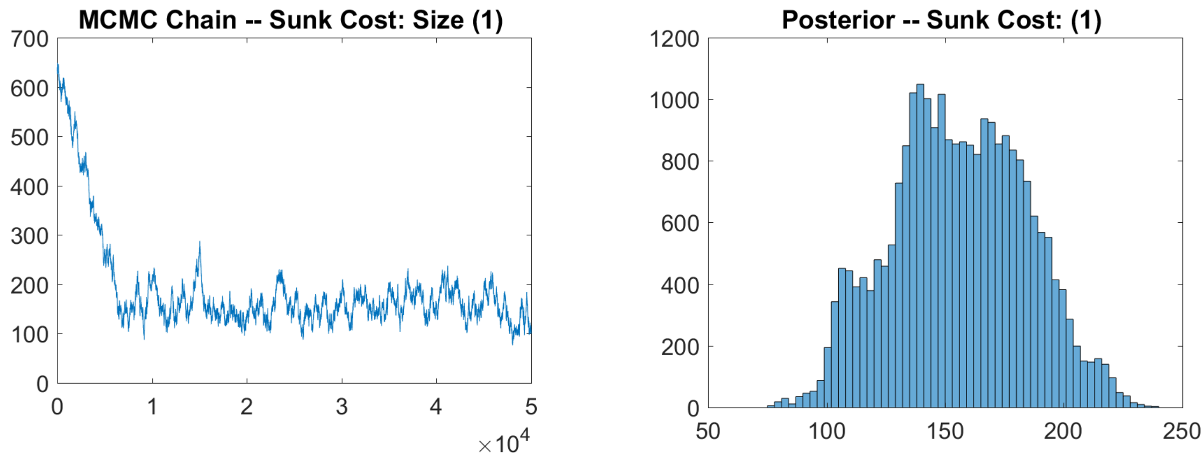
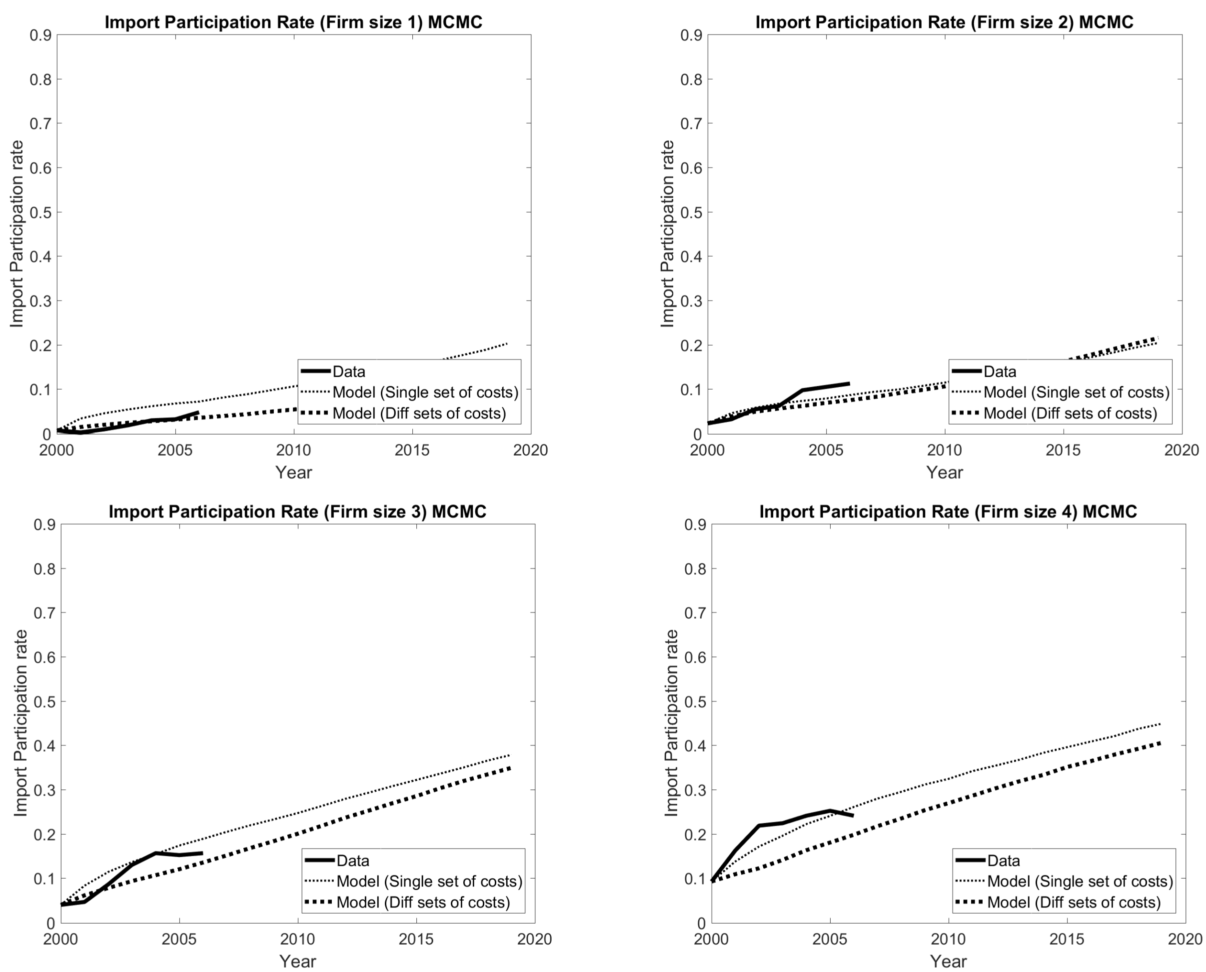


Figure C.4: Import Participation Rate by Firm Size



D Endogenize Intensive Margin of Import: Sketch

In this section, I will sketch the possible strategy to endogenize the intensive margin of import, which constitutes an important future task of this project. In doing so, I will measure N_{it} by the number of imported varieties the firm imports instead of using total intermediate input to imported intermediate input ratio. More precisely, N_{it} is one plus the number of HS products imported by firm i. N_{it} is normalized to one if a firm is not using imported input. Except for this modification, static component of the model is same as before.

Dynamic Import Decision

Firm's state vector $\mathbf{s}_{it}$ is same as before,

$$s_{it} = (\omega_{it}, d_{it-1}, k_{it}, \Phi_t, \mathbf{w}_{it})$$

which includes today's productivity ω_{it}, previous period import status d_{it-1}, predetermined capital stock k_{it}, the market aggregate variables Φ_t, and firm-time specific factor prices $\mathbf{w}_{it}$.

After observing state vector, a firm observes relevant cost of import and decides whether to import or not, and how much to import conditional on choosing to import. Making discrete import decision requires for a firm to compare the benefit and cost of import. The benefit consists of static and dynamic component. The former depends on the intensive margin of import (N_{it}). Therefore, I will introduce the fixed cost per variety to make the optimal choice of N_{it} as a function of state variables.

Given firm's state vector s_{it}, value function before observing fixed cost and sunk cost is, by choosing $d_{it} = \{0, 1\}$,

$$V(s_{it}) = \int \max_{d_{it}} \Big\{ u(d_{it}, s_{it}|\gamma_{it}) + \delta \mathbb{E}_t \left[V(s_{it+1}, d_{it}) \right] \Big\} dG^{\gamma}$$

where $u(d_{it}, s_{it}|\gamma_{it})$ is current period payoff conditional on relevant import cost γ_{it}, which can be written as

$$u(d_{it}, s_{it}|\gamma_{it}) = d_{it}^N \Pi_{it}^N + d_{it}^I \left[\Pi_{it}^I - (d_{it-1}^N \gamma_{it}^S + d_{it-1}^I \gamma_{it}^F)\right]$$

where operational profit before paying the fixed or sunk cost is import-decision-specific. If a firm chooses not to import, the current pay off function is same as before:

$$\Pi_{it}^N = \frac{1}{\sigma} R_{it}^N = \left[\frac{\sigma}{\sigma-1} MC_{it}\left(\omega_{it}, k_{it}, \mathbf{w}_{it}|d_{it}^N = 1\right)\right]^{1-\sigma} \frac{Y_t}{P_t^{1-\sigma}}$$

If a firm chooses to import, they have to pay the fixed cost per variety:

$$\begin{aligned}\Pi_{it}^I &= \max_{N_{it}} \frac{1}{\sigma} R_{it}^I - F(N_{it}) \\ &= \max_{N_{it}} \left[\frac{\sigma}{\sigma-1} MC_{it}\left(\omega_{it}, k_{it}, N_{it}, \mathbf{w}_{it}|d_{it}^N = 1\right)\right]^{1-\sigma} \frac{Y_t}{P_t^{1-\sigma}} - F(N_{it}) \qquad \text{(D.1)}\end{aligned}$$

$F(N_{it})$ is the cost to increase the number of imported variety. This is an increasing function of N_{it}, for example, $F(N_{it}) = N_{it}^\rho$ with $\rho > 0$. The RHS of equation (D.1) formally specifies the firm's optimization problem for the intensive margin of import. Since the policy function for the intensive margin of import can be written as $N_{it} = N(k_{it}, \omega_{it})$, importer's profit Π_{it}^I is a function of the state variables. Therefore, current period payoff (and value function) does not have N_{it} in its argument. The remaining component of the model is same as before.

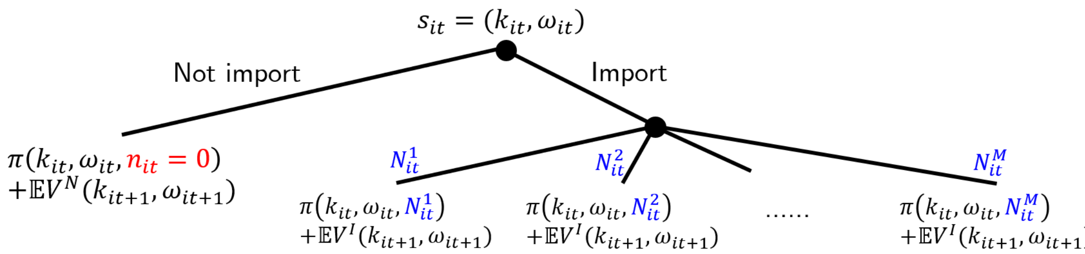
Empirical Strategy for the Second Stage

Only departure from the current manuscript is that I need to estimate the parameter of $F(N_{it})$, denoted by ρ. I will explain the algorithm to solve the value function and to construct the likelihood.

(1) Begin with an initial guess of parameters $\Theta^0 = (\lambda^{S0}, \lambda^{F0}, \rho^0)$ where $\lambda^m > 0$, $m = \{S, F\}$ is the exponential distribution parameter for sunk/fixed cost and $\rho > 0$ is parameter which governs the cost per variety.

(2) Calculate the transition probability of productivity $F(\omega'|\omega, d)$ based on estimated parameters in the first stage, which is used to calculate the expected continuation value below.

Figure D.1: Determining Intensive and Extensive Margin of Import



(3) Iterate on the following inner loop to find fixed points of value functions

(a) Begin with a set of initial guess of value functions $V^{m0}(s)$ for $m = \{N, I\}$. State variables are triplet, capital, productivity, and previous import status.

(b) Using the transition probability obtained in the previous stage, calculate the expected continuation value, where expectation is taken over the next period productivity.

$$\mathbb{E}V^{m0} = \int_{\omega'} V^{m0}(s')dF(\omega'|\omega, d^m) = \sum_h V^{m0}(k', d^m, \omega'_h) \Pr(\omega'_h|\omega, d^m)$$

(c) Update value functions for importer and non-importer according to equation (22) and (24). In computing importer's profit, I need to solve firm's optimization problem for the intensive margin of import. Figure 1 explains the structure of the problem. Suppose that N_{it} takes M possible values: $N_{it}^1, ...N_{it}^M$. First, given state variable (k_{it}, ω_{it}), compute current period pay off for each possible value of N_{it}, i.e., $\Pi(k_{it}, \omega_{it}, N_{it}^m)$ for $m = 1, .., M$. Then, pick m which gives the highest current payoff. This corresponds to the firm's optimization problem displayed in equation (D.1). This gives the optimal choice of intensive margin of import and in turn the current payoff of the firm as if it decided to import. Then we can compute the marginal benefit from importing according to equation (18), which gives the conditional choice probability.

(d) Iterate the step (a) – (c) until $|V^{m,n+1} - V^m| < \varepsilon$ where ε is predetermined tolerance level.

(4) Given the value function, evaluate the likelihood function of every observation.