Learning by Doing Model with Melitz (2003) which should have been “Learning by Exporting”

Wooyong Park
Yonsei University

Link to the Latest version

Last updated June 2, 2025

1 Motivation

- Motivating example: Liang et al. (2024)

Contents lists available at ScienceDirect

Journal of International Economics

journal homepage: www.elsevier.com/locate/jie

Full length articles

Learning by exporting: Evidence from patent citations in China☆

Yousha Liang[a], Kang Shi[b], Hanyi Tao[c,*], Juanyi Xu[d]

[a] School of Economics & Center for Research of Private Economy & Academy of Financial Research, Zhejiang University, Hangzhou, China
[b] Department of Economics, Chinese University of Hong Kong, Shatin, New Territories, Hong Kong
[c] School of Entrepreneurship and Management, ShanghaiTech University, Shanghai, China
[d] Department of Economics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

ARTICLE INFO

Dataset link: https://data.mendeley.com/datasets/m7mgyfdzwt/1

JEL classification:
F13
F14
O14
O31

Keywords:
Innovation
Patent citations
Knowledge transfers
Learning by exporting

ABSTRACT

Using matched firm-level data on operations, trade, and patents in China, we empirically investigate the impact of exporting on firms’ innovation and learning from destinations. We first show that access to export markets improves both the quantity and quality of innovation. We then construct a measure of knowledge flows based on Chinese patent citation data and demonstrate that there are actual knowledge transfers from destination markets to exporting firms, which help firms improve their innovation. These findings support the learning-by-exporting effect. Moreover, we show that this learning effect remains significant and unchanged, even after controlling for the spillover effect of FDI. To investigate how the learning process occurs, we also explore whether the learning-by-exporting effect varies with different export products or export modes. We find that vertical linkages and direct exporting facilitate Chinese firms’ learning from destination markets. Finally, we show that the learning effect is particularly salient at the extensive margin compared to the intensive margin.

- Melitz (2003) assumes that productivity determines export behavior.
- Empirical evidence that export also affects productivity
- How to accomodate the increasing productivity after export in Melitz (2003)

2 In this report I pin down

- LbE-FEC (learning by doing - free entry condition)
- LbE-ZCP (learning by doing - zero cutoff profit)

3 Set up

1. CES Demand as usual
2. Monopolistic competition at each variety ω
3. Firm technology:
$$l_{ij} = f_{ij} + \frac{\tau_{ij}}{\phi} q_{ij}$$
4. The distribution of productiviy ϕ follows
$$\begin{cases} \phi & \sim G(\gamma) \\ \phi^* & \sim K(\phi^* | \phi; \gamma^*) \quad (\phi^* > \phi) \end{cases} \tag{1}$$
where G and K are both Pareto distributions with different values of shape parameter: γ, γ^*

- ϕ^* is not observed by the firm even after entry. Rather, it draws again after the entry, because the firm's productivity changes after its production activity. ($\rightarrow$ Limitation: no explanation of the exact mechanism of learning)
- Melitz (2003) is a special case of (1) where $\gamma^* \rightarrow \infty$.

4 Zero-Cutoff Profit

In Melitz (2003), ZCP was determined by ϕ_0:

$$\pi(\phi_0) = 0$$

where

$$\pi_{ij}(\phi) = \frac{1}{\sigma} \left(\frac{\sigma}{\sigma - 1} w_i \tau_{ij} \right)^{1-\sigma} \mathbb{P}_j^{\sigma-1} Y_j \phi^{\sigma-1} - w_i f_{ij} \tag{2}$$

4.1 Key Difference

- Uncertainty still remains after entry, so ZCP is determined by the 'average' profit:

$$\mathbb{E}_{\phi^* \sim K}\left[\frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1}w_i\tau_{ij}\right)^{1-\sigma}\mathbb{P}^{\sigma-1}Y_j\phi^{*\sigma-1}\right] = w_i f_{ij}$$

- Let $\Psi \equiv \frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1}w_i\tau_{ij}\right)^{1-\sigma}\mathbb{P}^{\sigma-1}Y_j$. Then,

$$\begin{aligned}
w_i f_{ij} &= \Psi \cdot \mathbb{E}_{\phi^* \sim K}\left[\phi^{*\sigma-1}\right] \\
&= \Psi \cdot \int_{\phi}^{\infty} \phi^{*\sigma-1}\frac{\gamma^* \phi^{\gamma^*}}{\phi^{*\gamma^*+1}}d\phi^* \\
&= \Psi \cdot \int_{\phi}^{\infty} \frac{\phi^{\gamma^*-\sigma+1}}{\phi^{*\gamma^*-\sigma+2}} \quad \frac{\gamma^*}{\phi^{1-\sigma}}d\phi^* \\
&= \Psi \cdot \frac{\gamma^*}{\gamma^*-\sigma+1}\cdot\frac{1}{\phi^{1-\sigma}}\underbrace{\int_{\phi}^{\infty}(\gamma^*-\sigma+1)\frac{\phi^{\gamma^*-\sigma+1}}{\phi^{\gamma^*-\sigma+2}}d\phi^*}_{=1} \\
&= \Psi \cdot \frac{\gamma^*}{\gamma^*-\sigma+1}\cdot\frac{1}{\phi^{1-\sigma}}
\end{aligned}$$

- Thus, LbE-ZCP is determined by

$$\frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1}w_i\tau_{ij}\right)^{1-\sigma}\mathbb{P}^{\sigma-1}Y_j\frac{\gamma^*}{\gamma^*-\sigma+1}\cdot\phi^{\sigma-1} = w_i f_{ij} \tag{3}$$

- Compare this to the Melitz (2003) 's ZCP:

$$\frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1}w_i\tau_{ij}\right)^{1-\sigma}\mathbb{P}_j^{\sigma-1}Y_j\phi^{\sigma-1} = w_i f_{ij}$$

- Since $\frac{\gamma^*}{\gamma^*-\sigma+1} > 1$, the possibility of learning - and that firms expecting such learning - lowers the ZCP and more firms get to export.

4.2 Comparative statics

1. Increase of γ^* - **inverse: Learning Rate**

- γ^* is the measure of firm homogeneity. Because γ^* determines how firms shift from initial productivity ϕ to post-learing productivity ϕ^*, higher γ^* means less possibility of productivity jump.
- $\gamma^* \to \infty$: degenerate K and no learning, same as Melitz (2003)
- $\gamma^* \to \sigma - 1$: The first moment - mean - of K goes to ∞, every firm will export.

Proposition 4.1 (Learning and Export Behavior)**.** *If $\gamma^* \to \sigma - 1$, that is the learning rate reaches its maximum, all firms export regardless of its intial productivity level ϕ.*

5 Free Entry Condition (Sketch)

5.1 Melitz (2003)

How is FEC determined?

$$
\begin{aligned}
w_i f^e &= \mathbb{E}_{\phi \sim G}\Big[\sum_{j \in S} \max\{\pi_{ij}(\phi), 0\}\Big] \\
&= \int \sum_{j \in S} \max\{\pi_{ij}(\phi), 0\} dG_\gamma(\phi) \\
&= \int \sum_{j \in S} \max\{\pi_{ij}(\phi), 0\} \gamma \phi^{-\gamma-1} d\phi
\end{aligned} \tag{4}
$$

The LbE-FEC, the term $\max\{\pi_{ij}(\phi), 0\}$ is different from Melitz (2003) because the profit relies on ϕ^*, not ϕ. Therefore, LbE-FEC has to be summed over the joint density of ϕ^* and ϕ:

$$
w_i f^e = \int_\phi \int_{\phi^* > \phi} \sum_{j \in S} \max\{\pi_{ij}(\phi^*), 0\} \underbrace{\frac{\gamma^* \phi^{\gamma^*}}{\phi^{*\gamma^*+1}}}_{\text{cpdf of} \phi^*} d\phi^* \underbrace{\gamma \phi^{-\gamma-1}}_{\text{pdf of} \phi} d\phi \tag{5}
$$

We expand Equation 5 as follows:

$$
\int_\phi \Bigg(\underbrace{\int_0^\infty \sum_{j \in S} \max\{\pi_{ij}(\phi^*), 0\} \frac{\gamma^*}{\phi^{*\gamma^*+1}} d\phi^*}_{\text{similar to Melitz}} - \underbrace{\int_0^\phi \sum_{j \in S} \max\{\pi_{ij}(\phi^*), 0\} \frac{\gamma^*}{\phi^{*\gamma^*+1}} d\phi^*}_{\text{Suppose} \approx 0} \Bigg) \gamma \phi^{\gamma^* - \gamma - 1} d\phi
$$

"*Suppose*≈ 0" is a simplifying assumption.

Assume that the learning rate is not too large - i.e., γ^* has a lower bound, γ. Then, since $\gamma < \gamma^*$,

$$
\int \sum_{j \in S} \max\{\pi_{ij}(\phi), 0\} \gamma \phi^{-\gamma-1} d\phi > \int_0^\infty \sum_{j \in S} \max\{\pi_{ij}(\phi^*), 0\} \frac{\gamma^*}{\phi^{*\gamma^*+1}} d\phi^*
$$

and from Melitz (2003)

$$
(w_i f^e)_{\text{Melitz}} = \int \sum_{j \in S} \max\{\pi_{ij}(\phi), 0\} \gamma \phi^{-\gamma-1} d\phi
$$

Equation 5 becomes

$$
\begin{aligned}
\int_{\phi}\left(\int_0^\infty \sum_{j\in S}\max\{\pi_{ij}(\phi^*),0\}\frac{\gamma^*}{\phi^{*\gamma^*+1}}d\phi^*\right)\gamma\phi^{\gamma^*-\gamma-1}d\phi &< (w_i f^e)_{\text{Melitz}}\int_{\phi}\gamma\phi^{\gamma^*-\gamma-1}d\phi \\
&= (w_i f^e)_{\text{Melitz}}\frac{\gamma-\gamma^*}{\gamma}\int_{\phi}\frac{1}{\gamma-\gamma^*}\phi^{\gamma^*-\gamma-1}d\phi \\
&= (w_i f^e)_{\text{Melitz}}\left(1-\frac{\gamma^*}{\gamma}\right) \\
&< (w_i f^e)_{\text{Melitz}}
\end{aligned}
$$

Therefore, the cutoff for entry falls as well.

6 Remarks

- This model actually depicts how the expectation of learning affects entry decision, not how firms learn per se.
  - Distribution K can differ across individual firms, and $1/\gamma^*$ captures how firms are optimistic of future productivity increase.
  - Unless $\gamma^* \to \infty$, Melitz (2003) overestimates the cutoff prodictivities for ZCP and FEC.

7 Limitations

- Technically, this is "learning by doing", not "learning by exporting".
- It is also different from the lbe mechanism Liang et al. (2024) illustrates - f_{ij} decrease

8 Data Implications

8.1 Backing γ_i^* out

As mentioned above, distribution K can differ across individual firms, which imply $\gamma_i^* \neq \gamma_j^*$. In this Bayesian world, we can predict individual inverse learning rate γ_i^* by constructing posterior-prior.

8.1.1 Data Requirements

- $w_i, \sigma, \tau_{ij}, \mathbb{P}, Y_j, \phi, f_{ij}$
- ϕ, in my opinion, does not have to be observed but it does help pin down γ_i^*.
- If we do not have ϕ, having the export status data might help.

8.1.2 Set-up

- Prior

$$\gamma_i^* \sim H(\cdot) \quad \text{(prior)} \tag{6}$$

- Posterior is denoted by λ, and likelihood is denoted by L.

8.1.3 Results

$$\begin{aligned}\lambda(\gamma_i^* | \text{export status} = 1 | \gamma_i^*) &\propto L(\text{export status} = 1) \cdot H(\cdot) \\ &= Pr\left[\frac{1}{\sigma}\left(\frac{\sigma}{\sigma - 1} w_i \tau_{ij}\right)^{1-\sigma} \mathbb{P}^{\sigma-1} Y_j \frac{\gamma^*}{\gamma^* - \sigma + 1} \cdot \phi^{\sigma-1} > w_i f_{ij}\right] \cdot H(\cdot)\end{aligned}$$

- The prior, $H(\cdot)$, can be obtained by the usual Bayesian methods(e.g., strict prior, empirical bayes, heirachical bayes)
- The Bayes estimate for γ_i^* can be obtained as either the posterior mean or the mode(MaP estimate).

8.2 Using Athey et al. (2025)

Consider the following relation in an RCT:

$$Z \longrightarrow Y_1 \longrightarrow \underbrace{Y_2 = \phi^*}_{\text{unobs. in the RCT}}$$

- Z: Government support(that is targeted for productivity increase)
- Y_1: Immediately observable outcome of the RCT (e.g., test scores, patent application)
- Requirements:
  1. The RCT with (Z, Y_1)
  2. Observational (nonRCT) dataset with (Z, Y_1, Y_2) $\longrightarrow$ SIMS dataset
  3. Latent Unconfoundedness: Endogeneity pattern betweem (Z, Y_1) is the same as (Z, Y_2)

References

Athey, Susan, Raj Chetty, and Guido Imbens (2025) "The Experimental Selection Correction Estimator: Using Experiments to Remove Biases in Observational Estimates,"Technical report, National Bureau of Economic Research.

Liang, Yousha, Kang Shi, Hanyi Tao, and Juanyi Xu (2024) "Learning by exporting: Evidence from patent citations in China," 150, 103933.

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