Trade Liberalization, Export and Product Innovation

Sizhong Sun

College of Business Law and Governance
James Cook University
Queensland, Australia

Email: sizhong.sun@jcu.edu.au

[Abstract]

This paper studies firms' optimal response to a trade liberalization shock in terms of export and product innovation both theoretically and empirically. We find that trade liberalization, namely China's WTO accession, reduces trade cost and promotes export, which in turn incentivizes firms to innovate as the marginal benefit of innovation for exporting firms is higher than that for non-exporting firms. In addition, as a firm starts to innovate, it predicts to have a higher probability of moving to a better productivity state and can save the entry cost of innovation in the future, resulting in additional dynamic benefits. Such an innovation-promotion effect is an unintended consequence of trade liberalization.

[Key Words]

Trade Liberalization, Product Innovation, Export, Dynamic Discrete Choice Model, China

[JEL Classification]

F14, O31, L20, D22

* The author thanks Reshad Ahsan, Jasmine Jiang, Lingfei Lu, Phillip McCalman, Laura Puzzello and participants of the Online Australasian Seminar in International Economics, the 17th Australasian Trade Workshop, the 33rd Annual Conference of Chinese Economics Society Australia, the 2024 Econometric Society Australasia Meeting, and the 2025 Econometric Society Winter School in Dynamic Structural Econometrics for their comments and suggestions. The author also benefits from the 2022 Econometric Society Summer School in Dynamic Structural Econometrics. All errors remain the author's.

1. Introduction

Arguably, export and innovation play an important role in a firm's profitability and growth. Through export, firms expand into new markets, which may not only increase their profits but also bring in additional benefits, such as learning by exporting.[1] A successful product innovation boosts demand for a firm's products, which subsequently increases its profitability and promotes firm growth.

Given their importance, researchers have explored innovation and export from different dimensions, both theoretically and empirically.[2] One strand of the research is dedicated to investigating the bilateral interaction between export and innovation. On the one hand, international trade (export) can promote innovation. In a survey of the existing literature, Melitz and Redding (2021) highlight four key mechanisms through which international trade affects innovation, namely market size, competition, comparative advantage and knowledge spillovers. Empirical studies generally confirm a positive effect,[3] for example Salomon & Shaver (2005), Ayllón & Radicic (2019), De Fuentes, Niosi & Peerally (2021) and Lileeva & Trefler (2010).

On the other hand, innovation can boost firms' export, both in terms of extensive and intensive margins. For example, a successful process innovation augments productivity, which in turn promotes export. Empirical studies frequently confirm such a causality (see, among others, Becker & Egger, 2013; Cassiman, Golovko, & Martínez-Ros, 2010).

Many of these studies explore international trade and forms of innovation other than product innovation. In contrast, relatively fewer studies specifically focus on firms' export and product innovation.[4] Contributing to this strand of research, this paper investigates how a trade liberalization shock affects firms' optimal decision to export and how this impact is transmitted to the optimal decision of product innovation[5], both theoretically and empirically.

Theoretically, we model firms' behavior in a multi-stage decision process, where firms make optimal decisions on entry, exit, whether to innovate and export, the degree of product innovation, and prices in both the domestic and export markets. The model features mixed static and dynamic decision making. Decisions on optimal

---

[1] For example, see, among others, Bernard & Jensen (1999; 2004), Blalock & Gertler (2004), De Loecker (2007, 2013). Nevertheless, not all studies find evidence of learning by exporting, e.g. Clerides et al. (1998).

[2] To name a few, Egger and Keuschnigg (2015) on the role of finance in innovation and trade, Ederington and McCalman (2008) on the effects of international trade on technological diffusion, Atkeson and Burstein (2010) on how a change of the marginal cost of international trade affects firms' decisions to innovate and engage in international trade, and Impullitti and Licandro (2018) with oligopolistic competition.

[3] Researchers have also explored the impact of international trade on other forms of innovation (see for example Aghion, Bergeaud, Lequien, & Melitz, 2018; Bustos, 2011)

4 Empirical studies on the effect of export on product innovation include Damijan et al. (2010), Lin and Lin (2010), Bratti and Felice (2012), Olabisi (2017), Blyde et al. (2018), and Dai et al. (2020).

[5] Definition of product innovation can be found at OECD (2005).

degree of product innovation and prices in both markets are static. In contrast, decisions on entry, exit, and whether to innovate and export are dynamic in that they affect transition of firms' state variable.

Empirically, we bring the model to a sample of air-conditioner manufacturing firms in China from 2000 to 2007. Utilizing the shock of China's WTO accession in December 2001,[6] we quantify how it affects firms' export and how the effect is transmitted to firms' optimal decision to innovate. For identification, detailed in section 2.5, we impose the structure of theoretical model on the data, and associate different types of variation in the data with the identification of different parameters, similar to the identification strategy of Das, Roberts & Tybout (2007).

We find that China's WTO accession reduces the iceberg trade cost by around 13%, which unsurprisingly promotes export. The export promotion then further promotes innovation through three channels, one static and two dynamic. Statically, the contemporaneous marginal benefit of innovation for a given reduction of iceberg trade cost is larger for exporting firms, compared with non-exporting firms. Hence, it incentivizes exporting firms to innovate after trade liberalization, *ceteris paribus*. Dynamically, first, firms predict that export enables them to move to a better productivity state in the next period, which subsequently promotes the probability of innovation. Second, as a firm innovates, it can save entry costs of innovation in the future, which is substantial as per our estimation. This is a benefit of option value similar to that of export examined by Das et al. (2007), where we extend to product innovation.

In our estimations, we explicitly consider firms' decisions on entry and exit. If a firm decides not to enter or exit the market, it is not observed in the sample, resulting in an endogenous sample selection and attrition issue. Parameter estimates will be inconsistent and biased if the sample selection and attrition issues are not addressed properly. In this aspect, our study is distinct from the existing studies on product innovation and export.

The contribution of this study is two-fold, topic-wise and methodological. Topic wise, we document a mechanism how trade liberalization affects export and subsequently product innovation. That is, trade liberalization reduces the iceberg trade cost, *ceteris paribus*, which in turn promotes the value of export. The export promotion then improves firms' probability of product innovation through three channels, a contemporaneous static effect and two dynamic effects. Hence, firms are more likely to innovate after trade liberalization, *ceteris paribus*.

The contemporaneous effect corroborates with the findings of Bustos (2011), Lileeva & Trefler (2010), Atkeson & Burstein (2010), and Aw, Roberts & Xu (2011). The dynamic effect of transitioning to a better productivity state is similar to learning

[6] China's WTO accession is an overarching national initiative that is not likely to be driven by firm performance in the air-conditioning manufacturing industry, on which this study focuses. A number of existing studies have explored the impacts of China's WTO accession, for example Bloom et al. (2015), Lu and Yu (2015), Feng et al. (2017), Lin and Long (2020), Liu and Ma (2020), and Dai et al. (2020, 2021).

by exporting found by Bernard & Jensen (1999; 2004), Blalock & Gertler (2004), and De Loecker (2007, 2013), though in our context it is firms' prediction.[7] Besides, in addition to exporting, we also extend to the dimension of innovation. The dynamic benefit of option value corroborates with the finding of Das et al. (2007) regarding export, and we additionally extend to innovation. Compared with these studies, we explore these channels in a more integrated framework.

Methodologically, we establish an analytical framework which can be utilized in other similar setups. Our analytical framework features explicit consideration of firms' endogenous sample selection (entry) and attrition (exit). It is known that such sample selection and attrition lead to inconsistent and biased parameter estimates if not addressed properly. Existing studies focus more on the sample attrition issue (see for example Olley & Pakes, 1996), while the endogenous sample selection is less explored.[8] Sun (2023) considers both sample selection and attrition in estimating the effect of worker training on firms' labor productivity. However, his theoretical modelling is almost static in the sense that he focuses on the steady state. In contrast, we estimate a dynamic model of decisions on export and innovation, utilizing a nested fixed-point algorithm to compute the value functions, following Rust (1987).

The rest of the paper is organized into four sections. Section two presents the analytical framework, where we set up an economic model that features both static and dynamic decision making, and establish a four-step estimation algorithm. In section three, we discuss the data, variables, and reduced-form estimations. Section four reports the main estimation results, including the effects of trade liberalization and dynamic decisions on whether to innovate and export. We also conduct a counterfactual analysis in section four. Section five concludes the study.

2. Analytical Framework

2.1 Economic model

In each period, firms engage in a multi-stage decision making process, as described in Figure 1. Upon entering the market, randomly-sampled firms draw production (λ_1) and innovation (λ_2) capability endowments from an exogenous distribution with cumulative distribution function (CDF) of $G(\lambda_1, \lambda_2)$. At stage one, entry firms decide whether to enter the market by comparing the value of entry (the present value of all expected future profits) with an outside option. If they decide not to enter the market, they are not observed in the data, resulting in an endogenous sample selection.

At stage two, firms face shocks, either from the demand side (e.g., a decrease of consumers' disposable income) or from the supply side (e.g., an increase of input price), and decide whether to exit. If they decide to exit, they collect a liquidation value, normalized to 0, and are not observed in the data, leading to an endogenous sample attrition. Should they decide to stay, they then make decisions on whether to

[7] Rather than a causality.

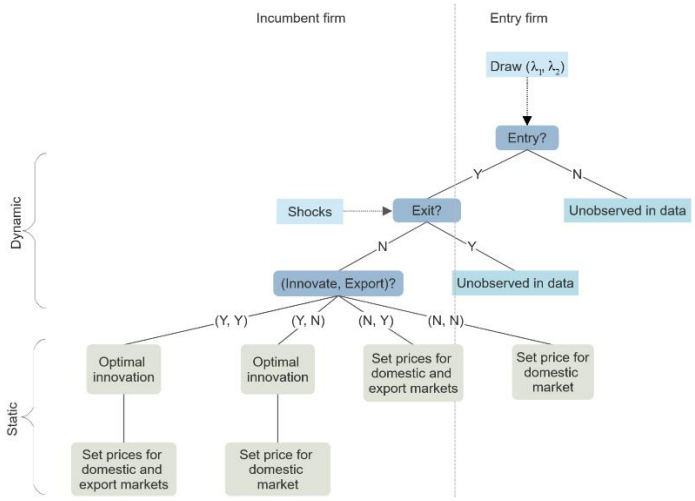
[8] Note such sample selection cannot be addressed by the classical Heckman sample selection model (Heckman, 1979) since firms are not observed in the sample if they decide not to enter the market. It is akin to truncation, which however is firms' endogenous decision (rather than exogenous truncation).

innovate and/or export at stage three. At stages four and five, firms choose an optimal level of product innovation and set prices for both domestic and export markets.[9]

Firms have perfect information in the current period, including the shocks at stage two. However, their information on the states in future periods is imperfect, and as such they form beliefs on the transition of the states (namely they predict the law of transition) when making dynamic decisions. As will be discussed in more details below, decisions at stages four and five are static, while decisions at stages one, two and three are dynamic.

The *equilibrium* is a set of optimal decisions on entry, exit, whether to innovate and/or export, optimal degree of product innovation, and prices in domestic and export markets, together with cut-off capabilities, that clear the markets and firms' beliefs are correct.[10]

Figure 1. The Decision Tree



The demand in the domestic market is $q = Ap^{1/\rho-1}$, where ρ is the constant-elasticity-of-substitution (CES) preference parameter ($0 < \rho < 1$); q is the quantity of goods that is a CES aggregation of new and existing products, namely $q \equiv \left[q_o^\rho + (\zeta q_n)^\rho\right]^{1/\rho}$ where the subscripts o and n denote existing and new products respectively,[11] we set the CES aggregation parameter to be ρ by choosing the numeraire appropriately, and ζ measures the attractiveness of new products, relative to existing products; A is aggregate demand, namely a representative consumer's

[9] Note our modelling features both ex ante and ex post firm heterogeneity.

[10] In the long run, the value of entry equals the entry cost.

[11] Note despite we do not frame our modelling in terms of multi-product firms, we are actually modelling multi-product firms with two types of products, the existing and new products. In addition, q_o and q_n can themselves be a CES aggregation of other multiple sub-types of existing and new products, respectively.

income divided by a price index; and p represents price.[12] Similarly, in the foreign market, the demand is $\tilde{q} = \tilde{A}\tilde{p}^{1/\tilde{\rho}-1}$, where tilde denotes the foreign market and $\tilde{q} \equiv \left[\tilde{q}_o^{\rho} + (\zeta \tilde{q}_n)^{\rho}\right]^{1/\rho}$.

Production incurs a fixed cost (*f*), due to facilities used in production. Using these facilities, firms can then combine one unit of labor with *m* units of intermediate inputs to produce *s* units of existing products. The term s measures labor productivity, and $s = \lambda_1 \psi$ where λ_1 and ψ are the unobserved and observed components of labor productivity respectively. Alternatively, to produce new products, firms pay an additional fixed cost of innovation (f_n), and combine one unit of labor with *m* units of intermediate inputs to produce $\lambda_2 s$ units of existing products. Let $\eta \equiv \frac{q_n}{q_o}$, which measures the level of product innovation. Then, the marginal cost of production can be written as $c = \frac{(\lambda_2+\eta)(w+m)}{\lambda_2(1+\zeta^{\rho}\eta^{\rho})^{1/\rho}s}$.

Firms' per-period profit from the domestic market is then $\pi = (p-c)q - f - \left[f_n + f_n^e\left(1-\chi_{1,-1}\right)\right]\chi_1$, where π denotes profit; χ_1 takes a value of one if the firm conducts product innovation and zero otherwise; $\chi_{1,-1}$ is the firm's product innovation status in the previous period; f_n and f_n^e are the fixed and entry costs of product innovation, respectively. If the firm has conducted innovation in the previous period, it does not need to pay the entry costs.[13] To export, firms need to pay a fixed and entry export costs (f_e and f_e^e, respectively) and an iceberg trade cost (τ). Export profit can then be written as $\tilde{\pi} = (\tilde{p} - \tau c)\tilde{q} - \left[f_e + f_e^e\left(1-\chi_{2,-1}\right)\right]$, where $\chi_{2,-1}$ is the firm's export status in the previous period; and similar to innovation, it saves on the entry cost if it has exported in the previous period.[14]

At stage five, firms set optimal prices as a markup over marginal cost of production, namely $p = \frac{c}{\rho}$ and $\tilde{p} = \frac{\tau c}{\tilde{\rho}}$ in the domestic and export markets, respectively. Accordingly, their per-period optimal profits are $\pi = (1 -$

---

[12] Such a demand function can be obtained by maximizing a CES utility function, subject to a budget constraint.

[13] More generally, as long as a firm innovates in the past, it can save the innovation entry cost, namely its innovation entry cost in year *t* is $f_n^e\left(1 - \mathbb{1}\left(\sum_{j=0}^{t}\chi_{1j} > 0\right)\right)$ where $\mathbb{1}(\cdot)$ denotes an indicator function taking a value of 1 if its argument is true and χ_{1j} takes a value of 1 if the firm innovates in year *j* and 0 otherwise.

[14] Similar to the innovation entry cost, one can specify a more general export entry cost in year *t* as $f_e^e\left(1 - \mathbb{1}\left(\sum_{j=0}^{t}\chi_{2j} > 0\right)\right)$ where χ_{2j} equals 1 if a firm exports in year *j* and 0 otherwise.

$\rho)\rho^{\frac{\rho}{1-\rho}}Ac^{\frac{\rho}{\rho-1}} - f - \left[f_n + f_n^e\left(1-\chi_{1,-1}\right)\right]\chi_1$ and $\tilde{\pi} = (1-\tilde{\rho})\rho^{\frac{\tilde{\rho}}{1-\tilde{\rho}}}\tilde{A}\tau^{\frac{\tilde{\rho}}{\tilde{\rho}-1}}c^{\frac{\tilde{\rho}}{\tilde{\rho}-1}} - \left[f_e - f_e^e\left(1-\chi_{2,-1}\right)\right]$. Hence, at stage four, firms' optimal level of product innovation, should they decide to innovate, is $\eta = \lambda_2^{\frac{1}{1-\rho}}\zeta^{\frac{\rho}{1-\rho}}$.

Therefore, firms' optimal per-period profits from domestic and export markets at stage four are $\pi(\chi_1 = 0) = (1-\rho)\rho^{\frac{\rho}{1-\rho}}A(w+m)^{\frac{\rho}{\rho-1}}\psi^{\frac{\rho}{1-\rho}}\lambda_1^{\frac{\rho}{1-\rho}} - f$, $\pi(\chi_1 = 1) = (1-\rho)\rho^{\frac{\rho}{1-\rho}}A(w+m)^{\frac{\rho}{\rho-1}}\psi^{\frac{\rho}{1-\rho}}\left(1+\lambda_2^{\frac{\rho}{1-\rho}}\zeta^{\frac{\rho}{1-\rho}}\right)\lambda_1^{\frac{\rho}{1-\rho}} - f - \left[f_n + f_n^e\left(1-\chi_{1,-1}\right)\right]$, $\tilde{\pi}(\chi_1 = 0) = (1-\tilde{\rho})\tilde{\rho}^{\frac{\tilde{\rho}}{1-\tilde{\rho}}}\tau^{\frac{\tilde{\rho}}{\tilde{\rho}-1}}\tilde{A}(w+m)^{\frac{\tilde{\rho}}{\tilde{\rho}-1}}\psi^{\frac{\tilde{\rho}}{1-\tilde{\rho}}}\lambda_1^{\frac{\tilde{\rho}}{1-\tilde{\rho}}} - \left[f_e + f_e^e\left(1-\chi_{2,-1}\right)\right]$, and $\tilde{\pi}(\chi_1 = 1) = (1-\tilde{\rho})\rho^{\frac{\tilde{\rho}}{1-\tilde{\rho}}}\tau^{\frac{\tilde{\rho}}{\tilde{\rho}-1}}\tilde{A}(w+m)^{\frac{\tilde{\rho}}{\tilde{\rho}-1}}\psi^{\frac{\tilde{\rho}}{1-\tilde{\rho}}}\left(1+\lambda_2^{\frac{\rho}{1-\rho}}\zeta^{\frac{\rho}{1-\rho}}\right)^{\frac{\tilde{\rho}}{1-\tilde{\rho}}\frac{1-\rho}{\rho}}\lambda_1^{\frac{\tilde{\rho}}{1-\tilde{\rho}}} - \left[f_e + f_e^e\left(1-\chi_{2,-1}\right)\right]$. Note $\frac{\partial\tilde{\pi}(\chi_1=1)}{\partial\tau} = \frac{\partial\tilde{\pi}(\chi_1=0)}{\partial\tau}\left(1+\lambda_2^{\frac{\rho}{1-\rho}}\zeta^{\frac{\rho}{1-\rho}}\right)^{\frac{\tilde{\rho}}{1-\tilde{\rho}}\frac{1-\rho}{\rho}} < \frac{\partial\tilde{\pi}(\chi_1=0)}{\partial\tau} < 0$. Hence, a decrease of the iceberg trade cost, for example due to trade liberalization, immediately incentivizes firms to innovate, should they start to export.

Let χ_2 denote whether a firm exports, taking a value of one if it does and zero otherwise. At stage three, if a firm decides to export (innovate), it needs to pay entry costs, which are waived if it has exported (innovated) in the previous period, due to re-utilizing facilities from the past period. In addition, decisions at stage three affect the evolution of state variable ($\frac{\psi}{w+m}$). Therefore, the stage three decisions are dynamic. In making decisions, firms form beliefs about the evolution of state variables. That is, they predict the law of transition.

Firms' *beliefs*. For the future shocks, which affect A and $\tilde{A}$, firms understand that there is a probability of $1-\sigma$ $(0<\sigma<1)$ that adversary shocks will occur, resulting in exit in the next period. Conditional on staying in the market, firms forecast their next period's A and $\tilde{A}$ by this period's A and $\tilde{A}$.[15] For the state variable ($\frac{\psi}{w+m}$), firms understand that it follows a finite-state Markov chain, conditional on entry and stay in the market, where the transition probability depends on today's innovation and export decisions. Firms predict the transition probability matrices based on their past experience and other available information. In equilibrium, firms' prediction of the law of transition is correct.

[15] Akin to an adaptive expectation.

Given firms' beliefs, we can obtain their value functions. The value of choosing $(\chi_1, \chi_2) = (0,0)$, normalized by the fixed cost of production, is $v_{(0,0)} \equiv \frac{1}{f}\breve{v}_{(0,0)} = (1-\rho)\rho^{\frac{\rho}{1-\rho}}\frac{1}{f}A(w+m)^{\frac{\rho}{\rho-1}}\psi^{\frac{\rho}{1-\rho}}\lambda_1^{\frac{\rho}{1-\rho}} - 1 + \delta\sigma E v_{(0,0)}$, where $\breve{v}$ and v denote the original and normalized value functions, respectively; δ is the discount rate and $Ev_{(0,0)}$ is the expected continuation value where the expectation is taken over the evolution of $\left(\frac{\psi}{w+m}\right)$, conditional on the choice of $(0,0)$. Similarly, the choice value of $(\chi_1, \chi_2) = (0,1)$ is $v_{(0,1)} \equiv \frac{1}{f}\breve{v}_{(0,1)} = (1-\rho)\rho^{\frac{\rho}{1-\rho}}\frac{1}{f}A(w+m)^{\frac{\rho}{\rho-1}}\psi^{\frac{\rho}{1-\rho}}\lambda_1^{\frac{\rho}{1-\rho}} + (1-\tilde{\rho})\tilde{\rho}^{\frac{\tilde{\rho}}{1-\tilde{\rho}}}\tau^{\frac{\tilde{\rho}}{\tilde{\rho}-1}}\frac{1}{f}\tilde{A}(w+m)^{\frac{\tilde{\rho}}{\tilde{\rho}-1}}\psi^{\frac{\tilde{\rho}}{1-\tilde{\rho}}}\lambda_1^{\frac{\tilde{\rho}}{1-\tilde{\rho}}} - 1 - \left[\frac{f_e}{f} + \frac{f_e^e}{f}\left(1-\chi_{2,-1}\right)\right] + \delta\sigma E v_{(0,1)}$. The choice value of $(\chi_1, \chi_2) = (1,0)$ is $v_{(1,0)} \equiv \frac{1}{f}\breve{v}_{(1,0)} = (1-\rho)\rho^{\frac{\rho}{1-\rho}}\frac{1}{f}A(w+m)^{\frac{\rho}{\rho-1}}\psi^{\frac{\rho}{1-\rho}}\left(1+\lambda_2^{\frac{\rho}{1-\rho}}\zeta^{\frac{\rho}{1-\rho}}\right)\lambda_1^{\frac{\rho}{1-\rho}} - 1 - \left[\frac{f_n}{f} + \frac{f_n^e}{f}\left(1-\chi_{1,-1}\right)\right] + \delta\sigma E v_{(1,0)}$. The choice value of $(\chi_1, \chi_2) = (1,1)$ is $v_{(1,1)} \equiv \frac{1}{f}\breve{v}_{(1,1)} = (1-\rho)\rho^{\frac{\rho}{1-\rho}}\frac{1}{f}A(w+m)^{\frac{\rho}{\rho-1}}\psi^{\frac{\rho}{1-\rho}}\left(1+\lambda_2^{\frac{\rho}{1-\rho}}\zeta^{\frac{\rho}{1-\rho}}\right)\lambda_1^{\frac{\rho}{1-\rho}} + (1-\tilde{\rho})\tilde{\rho}^{\frac{\tilde{\rho}}{1-\tilde{\rho}}}\tau^{\frac{\tilde{\rho}}{\tilde{\rho}-1}}\frac{1}{f}\tilde{A}(w+m)^{\frac{\tilde{\rho}}{\tilde{\rho}-1}}\psi^{\frac{\tilde{\rho}}{1-\tilde{\rho}}}\left(1+\lambda_2^{\frac{\rho}{1-\rho}}\zeta^{\frac{\rho}{1-\rho}}\right)^{\frac{\tilde{\rho}}{1-\tilde{\rho}}\frac{1-\rho}{\rho}}\lambda_1^{\frac{\tilde{\rho}}{1-\tilde{\rho}}} - 1 - \left[\frac{f_e}{f} + \frac{f_e^e}{f}\left(1-\chi_{2,-1}\right)\right] - \left[\frac{f_n}{f} + \frac{f_n^e}{f}\left(1-\chi_{1,-1}\right)\right] + \delta\sigma E v_{(1,1)}$. Note the normalization by fixed cost of production is without loss of generality and aids identification in the empirical estimations.

Accordingly, firms' optimal choices are as follows: $\chi_1 = \mathbb{1}\left(v_{(1,0)} > v_{(0,0)}\right)\mathbb{1}\left(v_{(1,0)} > v_{(0,1)}\right)\mathbb{1}\left(v_{(1,0)} \geq v_{(1,1)}\right) + \mathbb{1}\left(v_{(1,1)} > v_{(0,0)}\right)\mathbb{1}\left(v_{(1,1)} > v_{(0,1)}\right)\mathbb{1}\left(v_{(1,1)} > v_{(1,0)}\right)$, and $\chi_2 = \mathbb{1}\left(v_{(0,1)} > v_{(0,0)}\right)\mathbb{1}\left(v_{(0,1)} > v_{(1,0)}\right)\mathbb{1}\left(v_{(0,1)} \geq v_{(1,1)}\right) + \mathbb{1}\left(v_{(1,1)} > v_{(0,0)}\right)\mathbb{1}\left(v_{(1,1)} > v_{(1,0)}\right)\mathbb{1}\left(v_{(1,1)} > v_{(0,1)}\right)$, where $\mathbb{1}(\cdot)$ is the indicator function, taking a value of one if its argument is true and zero otherwise.

We note two observations here. First, since firms can save on entry costs if they export/innovate previously, they are more likely to export/innovate if they have done so previously. Second, as the per-period profits are monotone increasing functions of firm capability endowments (λ_1 and λ_2), we expect χ_1 and χ_2 to be positively correlated. That is, an exporting/innovating firm is more likely to innovate/export. Later we will examine the data regarding these two reduced-form predictions.

At stage two, if the realized adversary shock is substantially large, namely it reduces the per-period profit to a sufficiently large extent such that the value of staying in the market is less than its liquidation value, the firm exits. At stage one, firms' value functions depend on their capability endowments. If the realized capability endowments are such that the value functions are negative, firms do not enter the market. The stage two decision results in an endogenous sample attrition, while the stage one problem leads to an endogenous sample selection.

2.2 Estimation strategy

We intend to estimate the impacts of trade liberalization, namely China's WTO accession, on firms' export and product innovation. For this purpose, we implement the following four-step algorithm.

In step one, we first estimate the CES preference parameters, ρ and $\tilde{\rho}$, utilizing the structural relationship between total variable costs (TVC) and firms' domestic sales (R) and export revenues ($\tilde{R}$) implied by profit maximization at stage five, as follows:

$$TVC = \rho R + \tilde{\rho}\tilde{R} + \varepsilon_1 \quad (1)$$

where we append an error term ε_1 to capture the measurement error. Researchers have previously utilized a similar strategy to identify the CES preference parameters (see for example Aw et al., 2011).[16] Later in the estimations, we first regress total variable costs against domestic sales revenues, restricting the sample to non-exporting firms, to estimate ρ. Then using the estimated ρ, regress $(TVC - \rho R)$ against export revenues using the sample of exporting firms to estimate $\tilde{\rho}$.

In step two, we then proceed to estimate the impact of trade liberalization (WTO accession) on the iceberg trade cost. Firms' optimal per-period domestic sales and export revenues are $R = \rho^{\frac{\rho}{1-\rho}} A c^{\frac{\rho}{\rho-1}}$ and $\tilde{R} = \tilde{\rho}^{\frac{\tilde{\rho}}{1-\tilde{\rho}}} \tilde{A} c^{\frac{\tilde{\rho}}{\tilde{\rho}-1}} \tau^{\frac{\tilde{\rho}}{\tilde{\rho}-1}}$, where the aggregate demands A and $\tilde{A}$ are consumer income divided by an aggregate price index. In the estimation we will use the constant price gross national income (GNI) per capita divided by the base year price, in China and the world, to measure them.

Hence, $\frac{\tilde{\rho}\tilde{R}^{\frac{\tilde{\rho}-1}{\tilde{\rho}}} Y^{\frac{\rho-1}{\rho}}}{\rho R^{\frac{\rho-1}{\rho}} \tilde{Y}^{\frac{\tilde{\rho}-1}{\tilde{\rho}}}} = \frac{\tilde{P}^{\frac{\tilde{\rho}-1}{\tilde{\rho}}}}{P^{\frac{\rho-1}{\rho}}} \tau$, where Y and $\tilde{Y}$ are the constant price GNI per capita in China and the world, respectively; and P and $\tilde{P}$ are the base year prices in China and the world, respectively. Note with estimates of ρ and $\tilde{\rho}$ from step one and data of $(R, Y, \tilde{R}, \tilde{Y})$, the left-hand-side of the equation is observed. In light of the equation, we can use the following specification to identify the effect of WTO accession:

$$lny = \alpha_0 + \alpha_1 dWTO + \varepsilon_2 \quad (2)$$

[16] In Monte Carlo experiments, Sun and Anwar (2024) find that such a strategy has a high identification power.

where $lny = ln\left(\frac{\tilde{\rho}\tilde{R}^{\frac{\tilde{\rho}-1}{\tilde{\rho}}}Y^{\frac{\rho-1}{\rho}}}{\rho R^{\frac{\rho-1}{\rho}}\tilde{Y}^{\frac{\tilde{\rho}-1}{\tilde{\rho}}}}\right)$; $dWTO$ is a dummy variable, taking a value of one if post-WTO accession and zero otherwise;[17] ε_2 captures measurement errors and other firm level factors; $\alpha_0 = ln\left(\frac{\tilde{P}^{\frac{\tilde{\rho}-1}{\tilde{\rho}}}}{P^{\frac{\rho-1}{\rho}}}\right) + ln\tau_0$ where τ_0 is the pre-WTO accession iceberg trade cost; and $\alpha_1 = ln\tau_1 - ln\tau_0$ where τ_1 is the post-WTO accession iceberg trade cost.

The parameter α_1 measures change of iceberg trade cost after WTO accession. A similar identification strategy has been used in previous studies. For example, Alessandria and Choi (2014) identify the change of iceberg trade cost from change in the export-to-domestic sales ratio and the relative size of the markets. We expect trade liberalization to reduce the iceberg trade cost, namely a negative estimate of α_1. Besides, in equation (2), the calculation of lny uses the estimates of ρ and $\tilde{\rho}$ in step one. To account for the estimation errors brought in from step one, we compute the bootstrap standard errors in estimating equation (2) where we randomly re-sample (with placement) not just the data, but also from the distribution of the estimates of ρ and $\tilde{\rho}$ in step one.

In order to estimate the dynamic decisions on whether to innovate and export, we need to have the probability of exit and transition probability of state variable $\left(\frac{\psi}{w+m}\right)$. Since firms have correct belief (prediction) in equilibrium, we can nonparametrically estimate them as the relative frequency in step three. For the exit probability $(1-\sigma)$, it is the number of exit firms divided by the total number of firms in the sample, where a firm exits if it appears in the current year, but not in the next. That is, $\hat{\sigma} = \frac{1}{N}\sum_{i,t}\left(1 - \mathbb{1}(\Omega_{it})\right)$ where Ω_{it} is the event that firm i appears in year t but not in year $t+1$, the hat represents estimate, and N denotes the sample size. With an i.i.d. sample, $\hat{\sigma} \xrightarrow{p} \sigma$ and $\sqrt{N}(\hat{\sigma} - \sigma) \xrightarrow{d} \mathcal{N}[0, \sigma(1-\sigma)]$ as $N \to \infty$.

The state variable $\left(\frac{\psi}{w+m}\right)$ contains ψ and $(w+m)$, where $(w+m)$ equals total wage and intermediate inputs divided by the number of workers. For ψ, we first assume it depends on firm size (total number of workers) and capital intensity (fixed assets per worker). We then utilize the principal component analysis to extract the principal component from data of firm size and capital intensity. Dividing it by $(w+m)$, we further discretize it into *K* intervals. That is, $\left(\frac{\psi}{w+m}\right) \in \{1,2,\cdots,K\}$. Conditional on (χ_1, χ_2), the transition probability from one state to another is

[17] Note the WTO accession includes a mix of policies, such as tariff cut, that aim at trade liberalization. Existing studies use a same time dummy variable to measure WTO accession (for example Liu & Ma, 2020; Lu & Yu, 2015).

estimated as the share of firms that move from one state in the current year to another in the next year.

That is, conditional on (χ_1, χ_2), let $\mu_{j|k}$ denote the probability of transitioning from state k to state j, $\mu_{j,k}$ be the probability of being in state j currently and in state k previously, and μ_k be the probability in state k, where $k, j \in \{1,2,\cdots,K\}$. Then $\mu_{j|k} = \frac{\mu_{j,k}}{\mu_k}$. The probability $\mu_{j,k}$ can be estimated as $\hat{\mu}_{j,k} = \frac{1}{\tilde{N}}\sum_{i,t}\mathbb{1}\left(\left(\frac{\psi}{w+m}\right)_{it} = j, \left(\frac{\psi}{w+m}\right)_{it-1} = k\right)$, where $\tilde{N}$ denotes the size of sample with (χ_1, χ_2) and $\lim_{N\to\infty}\frac{\tilde{N}}{N} > 0$. Then $\hat{\mu}_{j,k} \xrightarrow{p} \mu_{j,k}$ as $N \to \infty$. Similarly, $\hat{\mu}_k = \frac{1}{\tilde{N}}\sum_{i,t}\mathbb{1}\left(\left(\frac{\psi}{w+m}\right)_{it-1} = k\right)$, and $\hat{\mu}_k \xrightarrow{p} \mu_k$ as $N \to \infty$. The transition probability is estimated as $\hat{\mu}_{j|k} = \frac{\hat{\mu}_{j,k}}{\hat{\mu}_k}$, where $\hat{\mu}_{j|k} \xrightarrow{p} \mu_{j|k}$ and the standard errors can be computed by the bootstrap method.

Step four estimates firms' decisions on whether to innovate and export, which is a dynamic discrete choice model.[18] For this purpose, we further parameterize $\hat{\lambda}_1 \equiv \frac{1}{f}A\lambda_1^{\frac{\rho}{1-\rho}} \sim e^{\mathcal{N}\left[-\frac{1}{2}ln2, ln2\right]}$, $\hat{\lambda}_2 \equiv \lambda_2^{\frac{\rho}{1-\rho}}\zeta^{\frac{\rho}{1-\rho}} \sim e^{\mathcal{N}\left[-\frac{1}{2}ln2, ln2\right]}$, $\tau^{\frac{\tilde{\rho}}{\tilde{\rho}-1}} f^{\frac{(1-\rho)\tilde{\rho}}{\rho(1-\tilde{\rho})}} \tilde{A} A^{-\frac{(1-\rho)\tilde{\rho}}{\rho(1-\tilde{\rho})}} \equiv e^{\beta_0}$, $\frac{f_e}{f} \equiv e^{\beta_1} + \varepsilon_3$, $\frac{f_n}{f} \equiv e^{\beta_2} + \varepsilon_4$, $\frac{f_e^e}{f} \equiv e^{\beta_3 + \beta_4 IsBig} + \varepsilon_5$, and $\frac{f_e^n}{f} \equiv e^{\beta_5 + \beta_6 IsBig} + \varepsilon_6$, where $\varepsilon_i \sim \mathcal{N}[0,1]$, $i = 3,4,5,6$; $IsBig \in \{0,1\}$ is a dummy variable, taking value 1 if a firm is big; and $\beta \equiv (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6)'$ is the vector of parameters that we will estimate from the data. Note $\hat{\lambda}_1$ and $\hat{\lambda}_2$ are log-normally distributed with mean and variance of 1.

Accordingly, the choice value functions become

$$v_{(0,0)} = (1-\rho)\rho^{\frac{\rho}{1-\rho}}\left(\frac{\psi}{w+m}\right)^{\frac{\rho}{1-\rho}}\hat{\lambda}_1 - 1 + \delta\sigma E v_{(0,0)},$$

$$v_{(0,1)} = (1-\rho)\rho^{\frac{\rho}{1-\rho}}\left(\frac{\psi}{w+m}\right)^{\frac{\rho}{1-\rho}}\hat{\lambda}_1 + (1-\tilde{\rho})\tilde{\rho}^{\frac{\tilde{\rho}}{1-\tilde{\rho}}}e^{\beta_0}\left(\frac{\psi}{w+m}\right)^{\frac{\tilde{\rho}}{1-\tilde{\rho}}}\hat{\lambda}_1^{\frac{(1-\rho)\tilde{\rho}}{\rho(1-\tilde{\rho})}} - 1 - \left\{e^{\beta_1} + \varepsilon_3 + \left(e^{\beta_3+\beta_4 IsBig} + \varepsilon_5\right)\left(1 - \chi_{2,-1}\right)\right\} + \delta\sigma E v_{(0,1)},$$

$$v_{(1,0)} = (1-\rho)\rho^{\frac{\rho}{1-\rho}}\left(\frac{\psi}{w+m}\right)^{\frac{\rho}{1-\rho}}\hat{\lambda}_1\left(1+\hat{\lambda}_2\right) - 1 - \left\{e^{\beta_2} + \varepsilon_4 + \left(e^{\beta_5+\beta_6 IsBig} + \varepsilon_6\right)\left(1 - \chi_{1,-1}\right)\right\} + \delta\sigma E v_{(1,0)}, \text{ and}$$

[18] Researchers have explored estimations of dynamic discrete choice models extensively, for example, to name a few, Hotz and Miller (1993), Imai et al. (2009), Keane and Sauer (2009), Keane and Wolpin (2009), Aguirregabiria (2010), Aguirregabiria and Mira (2010), Abbring (2010), Arcidiacono and Miller (2011), Bugni and Ura (2019), Abbring and Daljord (2020), and Kalouptsidi et al. (2021).

$$v_{(1,1)} = (1-\rho)\rho^{\frac{\rho}{1-\rho}}\left(\frac{\psi}{w+m}\right)^{\frac{\rho}{1-\rho}}\hat{\lambda}_1\left(1+\hat{\lambda}_2\right) + (1-$$

$$\tilde{\rho})\tilde{\rho}^{\frac{\tilde{\rho}}{1-\tilde{\rho}}}e^{\beta_0}\left(\frac{\psi}{w+m}\right)^{\frac{\tilde{\rho}}{1-\tilde{\rho}}}\hat{\lambda}_1^{\frac{(1-\rho)\tilde{\rho}}{\rho(1-\tilde{\rho})}}\left(1+\hat{\lambda}_2\right)^{\frac{(1-\rho)\tilde{\rho}}{\rho(1-\tilde{\rho})}} - 1 - \left\{e^{\beta_1} + \varepsilon_3 + \left(e^{\beta_3+\beta_4 IsBig} + \right.\right.$$

$$\left.\left.\varepsilon_5\right)\left(1-\chi_{2,-1}\right)\right\} - \left\{e^{\beta_2} + \varepsilon_4 + \left(e^{\beta_5+\beta_6 IsBig} + \varepsilon_6\right)\left(1-\chi_{1,-1}\right)\right\} + \delta\sigma E v_{(1,1)}.$$

The known parameters in the value functions, from estimations in the previous steps, are $(\rho, \tilde{\rho}, \sigma, \Sigma)$ where Σ represents the set of conditional transition probability matrices of discretized $\left(\frac{\psi}{w+m}\right)$, and we set the discount rate $(1-\delta)$ to 0.05.

In step four, first, we draw $D = 1000$ $\hat{\lambda}_1$ and $\hat{\lambda}_2$ from the standard log-normal distributions and $D = 1000$ ε_3, ε_4, ε_5 and ε_6 from the standard normal distributions.[19] Conditional on $(\rho, \tilde{\rho}, \sigma, \Sigma, \delta)$, one can compute the choice values for each pair of $\left(\hat{\lambda}_1, \hat{\lambda}_2\right)$ with a guess of β, namely $\left(v^d_{(0,0)}, v^d_{(0,1)}, v^d_{(1,0)}, v^d_{(1,1)}\right)$, $d = 1, \ldots, D$, where we utilize a nested fixed-point algorithm in the computation. Such a nested fixed-point algorithm has been utilized in computing value functions in other similar setups, pioneered by Rust (1987). Then, we can simulate the four choice probabilities, conditional on firm entry at stage one and stay at stage two. For example $\widehat{Pr}(\chi_1 = 1, \chi_2 = 1|entry, stay) = \frac{1}{\sum_{d=1}^{D}\mathbb{1}\left(v^d_{(0,0)}\geq 0\right)}\sum_{d=1}^{D}\frac{1}{1+exp\left(v^d_{(0,0)}-v^d_{(1,1)}\right)+exp\left(v^d_{(0,1)}-v^d_{(1,1)}\right)+exp\left(v^d_{(1,0)}-v^d_{(1,1)}\right)}$ where $\widehat{Pr}$ represents the McFadden kernel smoothed probability simulator. Note that since $\mathbb{1}\left(v^d_{(0,0)} \geq 0\right)$ does not depend on the parameter β, $\widehat{Pr}$ is a continuous function of β. In addition, by conditioning on firm entry and stay in the market, we address both the endogenous sample selection and attrition issues.

Since $E[\chi_1\chi_2|entry, stay] = Pr(\chi_{1it} = 1, \chi_{2it} = 1|entry, stay)$, a natural estimator is to choose parameters β to minimize the sum of squared errors between $\chi_1\chi_2$ and $\widehat{Pr}(\chi_1 = 1, \chi_2 = 1|entry, stay)$. That is,

$$\hat{\beta} = argmin\, Q \equiv \frac{1}{2N}\sum_{i,t}\left[\chi_{1it}\chi_{2it} - \widehat{Pr}(\chi_{1it} = 1, \chi_{2it} = 1|entry, stay)\right]^2 \qquad (3)$$

where $\hat{\beta}$ denotes the parameter estimate; N is the number of observations; and the subscripts (i, t) index firm-year observations. Later in our estimations, we first employ a simulated annealing algorithm to find initial values of the parameter estimate and then utilize the Nelder-Mead algorithm to find a solution to the minimization problem.[20] Besides, we use the bootstrap method to compute the standard errors.

---

[19] Note we implicitly assume these drawings are independent of each other.

[20] The estimation is implemented in R, using a *GenSA* package (Xiang, Gubian, Suomela, & Hoeng, 2013) for the simulated annealing algorithm and the in-built *optim* function for the Nelder-Mead algorithm.

In the estimations, the ***source of identification*** relies on particular types of variation in the data, a strategy similar to that of Das et al. (2007). The structural relationship between total variable costs and the domestic sales and export revenues allows identification of the preference parameters (ρ and $\tilde{\rho}$). The effect of WTO accession on iceberg trade cost is identified from the variation of the export-domestic-sales ratio from pre- to post-WTO accession. The identification of exit probability $(1-\sigma)$ utilizes time-variation of firms' disappearance in the sample, and similarly the transition probability matrices (Σ) are identified from time-variation of state variable $(\frac{\psi}{w+m})$.

The identification of sunk entry and fixed costs of export (β) in the dynamic choice modelling largely follows Das et al. (2007), and we extend it to product innovation. For export, the differences in exporting frequencies across firms, which only differ in terms of whether they have exported previously, allow identification of sunk entry costs. Then with the sunk entry costs, the frequency of turning between exporting and non-exporting among firms with positive gross profit streams identifies the fixed costs.

For product innovation, the sunk entry costs are identified from the differences in innovating frequencies among similar firms that only differ in terms of whether they have innovated in the previous period. Then, given the innovation entry cost and profits associated with product innovation, the frequency of turning between innovation and non-innovation identifies the fixed cost of innovation. Besides, as standard in the existing literature, we also impose restrictions on the discount rate (δ) and distributions of the shocks to aid identification.

One challenging task in the identification is to separate the static and dynamic effects. For example, given a static incentive to innovate, such as a decrease of trade costs due to WTO accession, firms that have never innovated may not be able to innovate immediately, owing to time needed to procure equipment and hire researchers. Hence, if the data contain only firms that have never innovated before the incentive shock, it may be difficult to separate the static and dynamic effects, as the static effect takes time to occur for these firms. In contrast, firms that have innovated previously can utilize the existing facilities to innovate, and thus can respond to the static incentive shock more quickly. Such variation in the responses allows for separation of the static and dynamic effects.

3. Data, Variables and Reduced-form Estimations

The data are a set of air-conditioner manufacturing firms in China, sourced from the National Bureau of Statistics (NBS), China. It covers 184 firms in 2000, 232 firms in 2001, 254 firms in 2002, 237 firms in 2003, 201 firms in 2005, 191 firms in 2006, and 206 firms in 2007, where years 2000 and 2001 are pre-WTO accession periods. It contains information on firms' domestic sales revenue (unit: thousand yuan), export revenue (unit: thousand yuan), value of new products (unit: thousand yuan), intermediate inputs (unit: thousand yuan), total wage (unit: thousand yuan), fixed assets annual net average (unit: thousand yuan), and the number of workers

(thousand). We deflate the monetary values by using the producer price index (base year: 2005) obtained from *China Statistical Yearbook 2008*.

Total variable cost (*TVC*) is the sum of total wage and intermediate inputs. We divide *TVC* by the number of workers to compute *wm* ($w + m$). Similarly, dividing the fixed assets annual net average by the number of workers, we can calculate the capital intensity (*kl*). Whether a firm conducts product innovation (χ_1) is a dummy variable that takes a value of one if it reports a positive new product value. Similarly, whether a firm exports (χ_2) is a dummy variable, taking a value of one if it reports positive export revenues.

Table 1. Summary Statistics

Variable	Obs	Mean	Std. dev.	Min	Max
R	1,505	423421.6	2127498	2.198614	3.33E+07
EX	1,505	160880.6	1032184	0	2.21E+07
TW	1,505	14575.22	78529.26	6.595843	2306854
M	1,505	464875.1	2312630	23	4.01E+07
TL	1,505	0.612936	2.195058	0.008	41
VNP	1,505	251432.1	2093122	0	4.17E+07
kl	1,505	114810.3	238748.2	0	3781611
wm	1,505	461799.2	595295.2	4600.29	7517665
χ_1	1,505	0.18804			
χ_2	1,505	0.352824			

Note: *R*: domestic sales revenues; *EX*: export revenues; *TW*: total wage; *M*: intermediate inputs; *TL*: the number of workers; *VNP*: value of new products; *kl*: fixed assets annual net average per worker; *wm*: *TVC* per worker; χ_1: whether a firm innovates; and χ_2: whether a firm exports.
Source: NBS, China, 2000-2007.

Table 1 reports the summary statistics. First, we observe substantial variation in the data. For example, the standard deviation for domestic sales revenues is as high as five times of its sample average. We will utilize these variations in the estimations. Second, more than 35% firm-year observations report export revenue, while less than 19% firm-year observations conduct product innovation. In the sample, 10% firm-year observations report both export and conduct product innovation. Third, on the intensive margin of product innovation, the average share of new product value in a firm's total sales revenue is 11.29%, and the export intensive margin is around the same level, with a sample average of 12.12%.

As illustrated in the theoretical modelling, we expect to observe a positive correlation between export and innovation and a positive correlation between export/innovation with their previous-period status. To check for these reduced-form predictions, we can nonparametrically estimate these probabilities from the sample as the relative frequency, similar to the estimate of transition probabilities.

Table 2 reports the estimates of probabilities, conditional on previous-period's status, which clearly confirm that firms that have innovated/exported in the previous

period are more likely to continue innovate/export. Compared with a previously non-exporting firm, a previously exporting firm has a probability of export that is three times higher. Similarly, a previously innovating firm has a probability of innovation that is more than six times higher than that of a previously non-innovating firm.

Table 2. Probabilities of Innovation and Export, conditional on Previous Period's Experience

	[1]		[2]	
	$\chi_1 = 0$	$\chi_1 = 1$	$\chi_2 = 0$	$\chi_2 = 1$
$\chi_{-1} = 0$	0.888 (0.0086)	0.112 (0.0086)	0.7808 (0.0119)	0.2192 (0.0119)
$\chi_{-1} = 1$	0.1988 (0.031)	0.8012 (0.031)	0.1213 (0.0187)	0.8787 (0.0187)

Note: For [1], the rows correspond to $\chi_{-1} = \chi_{1,-1}$; For [2], the rows correspond to $\chi_{-1} = \chi_{2,-1}$; Figures in the brackets are standard errors.

Table 3 reports the probabilities of innovation/export, conditional on contemporaneous export/innovation. Compared with a non-exporting firm, an exporting firm has a higher probability of conducting product innovation. Similarly, compared with a non-innovating firm, an innovating firm has a higher probability of export. That is, a positive correlation (complementarity) exists between the decisions on whether to innovate and export. The correlation between product innovation and export in the sample is 0.2034.

Table 3. Probabilities of Innovation/Export, conditional on Contemporaneous Export/Innovation

	[1] $Pr(\chi_1\|\chi_2)$			[2] $Pr(\chi_2\|\chi_1)$	
	$\chi_1 = 0$	$\chi_1 = 1$		$\chi_2 = 0$	$\chi_2 = 1$
$\chi_2 = 0$	0.8706 (0.0108)	0.1294 (0.0108)	$\chi_1 = 0$	0.6939 (0.0132)	0.3061 (0.0132)
$\chi_2 = 1$	0.7043 (0.0198)	0.2957 (0.0198)	$\chi_1 = 1$	0.4452 (0.0295)	0.5548 (0.0295)

Note: Figures in the brackets are standard errors.

4. Results

We estimate the theoretical model with a sample of air-conditioner manufacturing firms in China from 2000 to 2007, by implementing the algorithm described in Section 2.2.

In the first step estimation, we obtain the estimates of the CES preference parameters as $\hat{\rho} = 0.7517$ with a standard error of 0.0051 in the domestic market

and $\hat{\tilde{\rho}} = 0.9163$ with a standard error of 0.0079 in the export market, where the hats denote point estimates. The estimates are largely in line with findings in previous studies (see for example Aw et al., 2011; Das et al., 2007; Sun, 2023). Therefore, the elasticity of substitution in the domestic market is 4.03, lower than that in the export market (11.95). The export market consists of the world other than China, which is bigger and likely to have more options of air-conditioners for consumers. Hence, it is not surprising that consumers in the export market have a higher elasticity of substitution.

In step three, we nonparametrically estimate the probability of exit and the transition probability matrices. The exit probability is estimated as the share of firm-year observations where firms appear in this, but not next, year in the total firm-year observations. In the sample, 24.86% firm-year observations do not appear in the next year (standard error 0.0133). Conditional on innovation and export decisions (χ_1, χ_2), the transitional probability matrix of $\left(\frac{\psi}{w+m}\right)$ are nonparametrically estimated in a similar way as the relative frequency, where we restrict the sample to firms that appear in both adjacent years (namely conditional on firm entry and stay in the market).

4.1 The effect of WTO accession

Firms' profit maximization in the domestic and export markets implies a structural relationship between the domestic and export sales revenues, which depends on the CES preference parameters, the aggregate demands and iceberg trade cost. As shown in equation (2), with an appropriate measurement of the aggregate demands and the CES preference parameters, one can link a transformation of domestic sales and export revenues to the iceberg trade cost and a relative price in the base year. Since the base year relative price does not vary across time, one can compare the transformation of domestic sales and export revenues before and after WTO accession to identify the impact of trade liberalization.

The appended error term in equation (2) captures the measurement error and firm level factors. As a national initiative, WTO accession appears unlikely to be correlated with the measurement error and firm level factors in the air-conditioner manufacturing industry. Therefore, we estimate equation (2) by the random effect (RE) estimator. Table 4 reports the estimation results, where in column [1] *dWTO* = 1 if it is after 2001 and in column [2] *dWTO* is one year lagged to allow for the potential lag effect of WTO accession. The estimated coefficients in both columns are negative and statistically significant at the 5% level, suggesting that WTO accession significantly reduces firms' iceberg trade cost, which can occur due to a reduction of both tariff and non-tariff barriers. Specifically, compared with pre-WTO accession, the iceberg trade cost reduces by 13.5% after WTO accession (column [1]). The point estimate of the *dWTO* coefficient in column [2] is -0.184, higher than that of column [1]. Nevertheless, one estimate is within one standard deviation of the other.

To what degree do the estimations in columns [1] and [2] pick up the confounding effect of other unobserved factors?[21] For example, one may suspect that technological progress in the transportation sector can reduce the transportation cost, resulting in an estimate of a negative coefficient of *dWTO*. To check for this possibility, we estimate equation (2) with dummy variables for each year, rather than a *dWTO* dummy variable. If the effect picked up by estimate of α_1 is due to WTO accession, then we shall observe (1) the estimated coefficients of year dummies are not statistically significant in the pre-WTO periods, and significant in post-WTO periods; and (2) in post-WTO periods, the coefficients of year dummies shall be approximately constant across time.

Table 4. The Impact of Trade Liberalization on the Iceberg Trade Cost

	[1]	[2]	[3]
$\mathbb{1}(t = 2001)$			-0.0263
			(0.0877)
$\mathbb{1}(t = 2002)$			-0.0559
			(0.0907)
$\mathbb{1}(t = 2003)$			-0.186*
			(0.110)
$\mathbb{1}(t = 2005)$			-0.221**
			(0.0986)
$\mathbb{1}(t = 2006)$			-0.273***
			(0.100)
$\mathbb{1}(t = 2007)$			-0.266***
			(0.102)
constant	1.163***	1.170***	1.212***
	(0.150)	(0.153)	(0.155)
dWTO	-0.135***	-0.184***	
	(0.0496)	(0.0499)	
N	374	374	374

Note: In [1], *dWTO* = 1 if post-WTO accession (year > 2001); In [2], *dWTO* = 1 if year > 2002, namely one year lagged; $\mathbb{1}(\cdot)$ is the indicator function; Bootstrap standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: The authors' estimation with data from NBS, China.

Column [3] of Table 4 reports the results. One can observe that since 2003, the coefficients of dummy variables are significantly negative, which coincides with China's WTO accession in December of 2001. In addition, the estimates of the coefficients of dummy variables after 2003 are generally within one standard

[21] In particular, in equation (2), the left-hand-side variable has firm-year variations, while *dWTO* in the right-hand-side only has year variations.

deviation of each other, suggesting the reduction of iceberg trade cost does not significantly vary across years (after 2003). Hence, it appears that the reduction of iceberg trade cost is more driven by WTO accession, rather than other factors such as the technological progress in the transportation sector.[22]

4.2 The transition probability matrices

Table 5. The Transition Probability Matrixes

	1	2	3	4	1	2	3	4
	$(\chi_1 = 0, \chi_2 = 0)$				$(\chi_1 = 0, \chi_2 = 1)$			
1	0.7113	0.2535	0.0282	0.0070	0.7	0.2667	0.0333	0
	(0.0845)	(0.0468)	(0.0150)	(0.0075)	(0.1989)	(0.1089)	(0.0356)	(0)
2	0.2035	0.5752	0.1947	0.0265	0.1136	0.5909	0.2955	0
	(0.0425)	(0.0764)	(0.0415)	(0.0148)	(0.0545)	(0.1404)	(0.0924)	(0)
3	0.0194	0.1748	0.7184	0.0874	0.0227	0.1136	0.7045	0.1591
	(0.0134)	(0.0417)	(0.0940)	(0.0289)	(0.0208)	(0.0476)	(0.1345)	(0.0570)
4	0.0175	0.0526	0.1579	0.7719	0	0.0202	0.0808	0.8990
	(0.0177)	(0.0310)	(0.0556)	(0.1453)	(0)	(0.0145)	(0.0291)	(0.1017)
	$(\chi_1 = 1, \chi_2 = 0)$				$(\chi_1 = 1, \chi_2 = 1)$			
1	0.6429	0.3571	0	0	0.5	0.25	0	0.25
	(0.2923)	(0.2020)	(0)	(0)	(0.4943)	(0.3205)	(0)	(0.3205)
2	0	0.6190	0.3333	0.0476	0	0.6667	0.3333	0
	(0)	(0.2007)	(0.1392)	(0.0494)	(0)	(0.3589)	(0.2303)	(0)
3	0.0417	0.0833	0.6667	0.2083	0.0323	0.0323	0.8065	0.1290
	(0.0403)	(0.0574)	(0.1770)	(0.0926)	(0.0335)	(0.0335)	(0.1883)	(0.0682)
4	0.0526	0	0.1579	0.7895	0	0	0.0227	0.9773
	(0.0507)	(0)	(0.0898)	(0.2265)	(0)	(0)	(0.0217)	(0.1362)

Note: Transition is from rows to columns; For example, with $(\chi_1 = 0, \chi_2 = 0)$, a firm at state 1 has a probability of 0.2535 to transition to state 2 in the next period; Figures in brackets are standard errors, and note that when a probability is estimated to be 0/1, its S.E. is 0.

Source: The authors' estimation with data from NBS, China.

Table 5 reports the non-parametric estimates of the transition probability matrices ($\hat{\Sigma}$) of state variable $\left(\frac{\psi}{w+m}\right)$, conditional on (χ_1, χ_2). The estimates suggest that firms predict that export and innovation help them move to a better state in the next period, namely $\hat{\Sigma}(\chi_1 = 1, \chi_2 = 1)$, $\hat{\Sigma}(\chi_1 = 1, \chi_2 = 0)$ and $\hat{\Sigma}(\chi_1 = 0, \chi_2 = 1)$ first-order stochastically dominate $\hat{\Sigma}(\chi_1 = 0, \chi_2 = 0)$.[23] For example, at state 1, a firm has a

[22] We caution readers that observing the pattern in column [3] does not imply the estimate of α_1 picks up the effect of WTO accession. It is a necessary, but not sufficient, condition for our claim.

[23] Note the transition probability matrices represent firms' prediction of the law of transition, and only capture the extensive margin of innovation, it is unclear whether product innovation exhibits

probability of 0.2887 moving to a better state in the next period if it does not innovate or export. In contrast, this probability is 0.3 if it exports, 0.3571 if it innovates, and 0.5 if it does both.

4.3 The dynamic decisions on export and product innovation

As discussed in Section 2.2, we implement a nested fixed-point algorithm in estimating the dynamic choices of export and product innovation. We utilize the post-WTO data in the estimation, leaving the pre-WTO data for counterfactual analysis.

Table 6. Estimation Results of the Dynamic Choices of Export and Product Innovation

		Coef.	S.E.	t
Constant	β_0	-8.9705***	0.1367	-65.6280
Fixed cost of export	β_1	0.4345	0.3621	1.2001
Fixed cost of innovation	β_2	-1.7658***	0.5323	-3.3174
Entry cost of export	β_3	8.3171***	0.2412	34.4889
Entry cost of export, big firm	β_4	-2.3615***	0.3556	-6.6407
Entry cost of innovation	β_5	8.4302***	0.2693	31.2991
Entry cost of innovation, big firm	β_6	-1.404***	0.4116	-3.4114

Note: Standard errors are bootstrapped; *** p<0.01. Fixed and entry costs are relative to fixed cost of production.

Source: The authors' estimation with data from NBS, China.

Table 6 reports the estimation results. The parameter β_0, which contains the post-WTO iceberg trade cost ($\tau^{\frac{\tilde{\rho}}{\tilde{\rho}-1}} f^{\frac{(1-\rho)\tilde{\rho}}{\rho(1-\tilde{\rho})}} \tilde{A} A^{-\frac{(1-\rho)\tilde{\rho}}{\rho(1-\tilde{\rho})}} \equiv e^{\beta_0}$), is estimated to be close to -8.9 (namely $e^{\beta_0} = 0.000127$), statistically significant at the 5% level. Hence, with a 13.5% decrease of iceberg trade cost due to the WTO accession, the pre-WTO β_0 is -10.44926.

Regarding the fixed costs of export ($\frac{f_e}{f} \equiv e^{\beta_1} + \varepsilon_3$), the point estimate of β_1 is positive, but not statistically significant at the 5% level. The point estimate of β_2 is negative and significant at the 5% level, suggesting a relatively small fixed cost of innovation ($\frac{f_n}{f} \equiv e^{\beta_2} + \varepsilon_4$). In contrast, the entry costs of export and innovation appear to be large. For the export entry cost ($\frac{f_e^e}{f} \equiv e^{\beta_3 + \beta_4 IsBig} + \varepsilon_5$), the point estimates of β_3 and β_4 are 8.3171 and -2.3615, respectively, both of which are significant at the 5% level. Similarly, the point estimates of β_5 and β_6 ($\frac{f_e^n}{f} \equiv e^{\beta_5 + \beta_6 IsBig} + \varepsilon_6$) are 8.4302 and -1.404, respectively and significant at the 5% level.

increasing/decreasing returns. However, Bilir and Morales (2020) find decreasing returns to innovation.

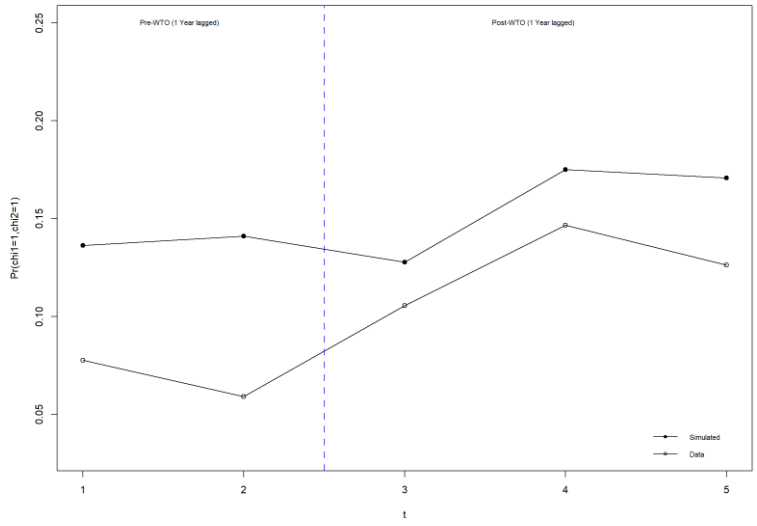
The pattern of small fixed costs and large entry costs corroborates with the finding of Das et al. (2007). The large entry costs of export and innovation point at an additional dynamic benefit from trade liberalization, namely as a firm innovates (and exports) in the current period, they can save the entry costs, which are substantial, in the future.

4.4 Counterfactual analysis

With the estimates above, we can then conduct counterfactual analysis to assess the effects of trade liberalization (WTO accession) on firms' probability to export and innovate. First, from the data, we can nonparametrically calculate this probability as the relative frequency of firms that both export and innovate in each year, namely $\widetilde{\Pr}(\chi_{1t}=1,\chi_{2t}=1)=\frac{1}{N_t}\sum_{i=1}^{N_t}\mathbb{1}(\chi_{1it}=1,\chi_{2it}=1)$ where $\widetilde{\Pr}$ denotes the probability estimate; and N_t is the number of firms in year *t*.

Let $z\equiv\left(\frac{\psi}{w+m},\chi_{1,-1},\chi_{2,-1},IsBig\right)$ and $Z\equiv\{z\}$, where $\frac{\psi}{w+m}$ is the discretised states and $|Z|=32$. Second, using the estimated parameters, we can simulate $\widetilde{\Pr}(\chi_{1t}=1,\chi_{2t}=1|z)$ where $\widetilde{\Pr}$ represents the simulated probability. Then $\widetilde{\Pr}(\chi_{1t}=1,\chi_{2t}=1)=\sum_{z_t\in Z}\widetilde{\Pr}(\chi_{1t}=1,\chi_{2t}=1|z)\widehat{\Pr}(z_t)$, where $\widehat{\Pr}(z_t)$ is computed nonparametrically from data as the relative frequency in each year. The probability $\widetilde{\Pr}(\chi_{1t}=1,\chi_{2t}=1)$ serves as the counterfactual. We exclude the samples of years 2000 and 2005 due to missing $(\chi_{1,-1},\chi_{2,-1})$, and the post-WTO is after 2002 (one year after the WTO accession) to allow for lagged effect from WTO accession.

Figure 2. The Probabilities of Product Innovation and Export



Note: *t* = 1, year 2001; *t* = 2, year 2002; *t* = 3, year 2003; *t* = 4, year 2006; *t* = 5, year 2007.

For post-WTO accession, the difference between $\widetilde{\Pr}(\chi_{1t}=1,\chi_{2t}=1)$ and $\widetilde{\Pr}(\chi_{1t}=1,\chi_{2t}=1)$ is due to estimation errors. For pre-WTO accession, this

difference is due to both estimation errors and a higher iceberg trade cost. Hence, we can infer the effect of WTO accession on these probabilities in a difference-in-differences manner.

Figure 2 presents the results. In all five years, the simulated probabilities of export and product innovation appear to be higher than their non-parametrically computed counterparts. However, the difference after WTO accession is smaller than that pre-WTO accession. Then the difference between these two differences is due to WTO accession that lowers the iceberg trade cost. Should there be a lower iceberg trade cost pre-WTO, the probability of export and product innovation will be higher than overserved in the data. That is, the WTO accession indeed promotes firms' probability of export and product innovation.

4.5 Robustness

A sizable proportion of exporting firms in the air-conditioner manufacturing industry is engaged in processing trade, where a firm imports all or part of the raw and auxiliary materials, parts, components, and packaging materials, and then re-exports the finished products after processing or assembly. Conceptually, it is likely that processing trade can affect a firm's product innovation. For example, via processing trade, a firm can learn how to manufacture a new product, which it then engineers for domestic market. To avoid such a possible confounding effect, we re-estimate the dynamic decisions on export and product innovation, excluding the processing trade firms.

Table 7. Dynamic Choice Model Estimation Results with Non-processing Trade Firms

		Coef.	S.E.	t
Constant	β_0	-10.0121***	0.2511	-39.8734
Fixed cost of export	β_1	0.5266	0.6547	0.8043
Fixed cost of innovation	β_2	-1.0505	1.0264	-1.0235
Entry cost of export	β_3	8.3742***	0.7051	11.8767
Entry cost of export, big firm	β_4	-1.9886 ***	0.9386	-2.1186
Entry cost of innovation	β_5	7.1497***	0.6797	10.5195
Entry cost of innovation, big firm	β_6	-1.6365	0.9346	-1.751

Note: Standard errors are bootstrapped; *** p<0.01. Fixed and entry costs are relative to fixed cost of production.

Source: The authors' estimation with data from NBS, China.

To do so, we first check the Customs data to see whether an exporting firm was engaged in processing trade. Firms are classified as processing trade firms if they participated in processing trade in at least one year during 2000, 2001, 2002, 2003, 2005 and 2006, which accounts for around 51% of the exporting firms. Then, we restrict the sample to non-processing trade firms, and re-estimate the dynamic choice model of export and product innovation. Table 7 reports the estimation results. Compared with Table 6, we can observe that, despite of difference in the magnitude

of point estimates, the estimation results are consistent with those of Table 6. In particular, the pattern of large entry costs and small fixed costs continues to hold. Therefore, our finding is robust to the potential confounding effect of processing trade.

Table 8. Two-state Transition Probability Matrixes

	1	2	1	2
	$(\chi_1 = 0, \chi_2 = 0)$		$(\chi_1 = 0, \chi_2 = 1)$	
1	0.8057	0.1943	0.7241	0.2759
	(0.0271)	(0.0271)	(0.0648)	(0.0648)
2	0.1324	0.8676	0.0503	0.9497
	(0.023)	(0.023)	(0.0188)	(0.0188)
	$(\chi_1 = 1, \chi_2 = 0)$		$(\chi_1 = 1, \chi_2 = 1)$	
1	0.7037	0.2963	0.7143	0.2857
	(0.0796)	(0.0796)	(0.1569)	(0.1569)
2	0.098	0.902	0.037	0.963
	(0.0432)	(0.0432)	(0.025)	(0.025)

Source: The authors' estimation with data from NBS, China.

Table 9. Two-state Estimation Results

		Coef.	S.E.	t
Constant	β_0	-5.3221***	0.365	-14.5808
Fixed cost of export	β_1	-4.5298	2.5625	-1.7677
Fixed cost of innovation	β_2	-4.8179***	2.3083	-2.0872
Entry cost of export	β_3	5.8097***	0.5763	10.0801
Entry cost of export, big firm	β_4	-1.9328***	0.814	-2.3745
Entry cost of innovation	β_5	7.5651***	0.3785	19.9847
Entry cost of innovation, big firm	β_6	-0.2653	0.3842	-0.6906

Note: Standard errors are bootstrapped; *** p<0.01. Fixed and entry costs are relative to fixed cost of production.

Source: The authors' estimation with data from NBS, China.

An empirical pattern of export and innovation is that only a small proportion of firms exports or innovates, and an even smaller proportion of firms does both. Consequently, in estimating the transition probability matrices, particularly that of $(\chi_1 = 1, \chi_2 = 1)$, an estimate of zero probability can be due to either the true transition probability being zero or lack of data. To check whether our findings are robust here, we also conduct estimations with two states. Table 8 reports the estimates of two-state transition probability matrices, and Table 9 presents the corresponding estimation results of the dynamic choice model of export and product innovation. Comparing Tables 8 and 9 with Tables 5 and 6, first, not surprisingly, we can observe differences in point estimates. Second, our findings continue to hold with two states. That is, firms predict that export and innovation help them move to a better state in the next period, and the entry costs of both export and innovation are large, while

fixed costs are small. Hence, our findings appear to be robust to different levels of states.

5. Concluding Remarks

This study explores firms' optimal decisions to export and conduct product innovation both theoretically and empirically. Theoretically, we model firms' decisions on entry, exit, whether to innovate and export, the degree of product innovation, and price setting in both the domestic and export markets. The theoretical model features a mixed static and dynamic decision-making process, and explicitly considers firms' entry and exit decisions, which result in endogenous sample selection and attrition issues for empirical estimations. Under the guidance of theoretical model, we then establish a four-step estimation algorithm, where we address both the endogenous sample selection and attrition issues.

Empirically, we bring the model to a sample of air-conditioner manufacturing firms in China from 2000 to 2007, in order to estimate the effect of China's WTO accession in 2001. Such a trade liberalization shock allows us to estimate its effect on export and subsequently how the effect is transmitted to product innovation. We find that China's WTO accession reduces the iceberg trade cost and unsurprisingly promotes export. In turn, the export promotion improves firms' product innovation through one static channel and two dynamic channels.

Statically, trade liberalization promotes product innovation through a larger contemporaneous marginal benefit of innovation from a given reduction of iceberg trade cost for exporting firms, which incentivizes them to innovate. Dynamically, export and product innovation helps firms to predict a higher probability of transition to a better state in the future. In addition, as firms export and innovate, they can save on the entry costs, which are substantial as per estimations, in the future, namely a benefit of option value.

Trade liberalization is generally considered to be beneficial in the sense of promoting international trade, particularly for developing economies. In addition, our findings suggest that it can also promote innovation, both in the short and longer terms. Such an additional benefit is to some extent an intended consequence of trade liberalization.

References

Abbring, J. H. (2010). Identification of Dynamic Discrete Choice Models. *Annual Review of Economics, 2*(Volume 2, 2010), 367-394. doi:https://doi.org/10.1146/annurev.economics.102308.124349

Abbring, J. H., & Daljord, Ø. (2020). Identifying the discount factor in dynamic discrete choice models. *Quantitative Economics, 11*(2), 471-501. doi:https://doi.org/10.3982/QE1352

Aghion, P., Bergeaud, A., Lequien, M., & Melitz, M. J. (2018). The Heterogeneous Impact of Market Size on Innovation: Evidence from French Firm-Level Exports. *National Bureau of Economic Research Working Paper Series, No. 24600*. doi:10.3386/w24600

Aguirregabiria, V. (2010). Another Look at the Identification of Dynamic Discrete Decision Processes: An Application to Retirement Behavior. *Journal of Business & Economic Statistics, 28*(2), 201-218. doi:10.1198/jbes.2009.07072

Aguirregabiria, V., & Mira, P. (2010). Dynamic discrete choice structural models: A survey. *Journal of Econometrics, 156*(1), 38-67. doi:https://doi.org/10.1016/j.jeconom.2009.09.007

Alessandria, G., & Choi, H. (2014). Do falling iceberg costs explain recent U.S. export growth? *Journal of International Economics, 94*(2), 311-325. doi:https://doi.org/10.1016/j.jinteco.2014.08.009

Arcidiacono, P., & Miller, R. A. (2011). Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity. *Econometrica, 79*(6), 1823-1867. doi:https://doi.org/10.3982/ECTA7743

Atkeson, A., & Burstein, A. (2010). Innovation, Firm Dynamics, and International Trade. *Journal of Political Economy, 118*(3), 433-484. doi:10.1086/653690

Aw, B. Y., Roberts, M. J., & Xu, D. Y. (2011). R&D Investment, Exporting, and Productivity Dynamics. *American Economic Review, 101*(4), 1312-1344. doi:doi: 10.1257/aer.101.4.1312

Ayllón, S., & Radicic, D. (2019). Product innovation, process innovation and export propensity: persistence, complementarities and feedback effects in Spanish firms. *Applied Economics, 51*(33), 3650-3664. doi:10.1080/00036846.2019.1584376

Becker, S. O., & Egger, P. H. (2013). Endogenous product versus process innovation and a firm's propensity to export. *Empirical Economics, 44*(1), 329-354. doi:10.1007/s00181-009-0322-6

Bernard, A. B., & Jensen, B. J. (1999). Exceptional exporter performance: cause, effect, or both? *Journal of International Economics, 47*(1), 1-25. doi:https://doi.org/10.1016/S0022-1996(98)00027-0

Bernard, A. B., & Jensen, J. B. (2004). Why Some Firms Export. *The Review of Economics and Statistics, 86*(2), 561-569. Retrieved from http://www.jstor.org/stable/3211647

Bilir, L. K., & Morales, E. (2020). Innovation in the Global Firm. *Journal of Political Economy, 128*(4), 1566-1625. doi:10.1086/705418

Blalock, G., & Gertler, P. J. (2004). Learning from exporting revisited in a less developed setting. *Journal of Development Economics, 75*(2), 397-416. doi:https://doi.org/10.1016/j.jdeveco.2004.06.004

Bloom, N., Draca, M., & Van Reenen, J. (2015). Trade Induced Technical Change? The Impact of Chinese Imports on Innovation, IT and Productivity. *The Review of Economic Studies, 83*(1), 87-117. doi:10.1093/restud/rdv039

Blyde, J., Iberti, G., & Mussini, M. (2018). When does innovation matter for exporting? *Empirical Economics, 54*(4), 1653-1671. doi:10.1007/s00181-017-1274-x

Bratti, M., & Felice, G. (2012). Are Exporters More Likely to Introduce Product Innovations? *The World Economy, 35*(11), 1559-1598. doi:https://doi.org/10.1111/j.1467-9701.2012.01453.x

Bugni, F. A., & Ura, T. (2019). Inference in dynamic discrete choice problems under local misspecification. *Quantitative Economics, 10*(1), 67-103. doi:https://doi.org/10.3982/QE917

Bustos, P. (2011). Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms. *American Economic Review, 101*(1), 304-340. doi:10.1257/aer.101.1.304

Cassiman, B., Golovko, E., & Martínez-Ros, E. (2010). Innovation, exports and productivity. *International Journal of Industrial Organization, 28*(4), 372-376. doi:https://doi.org/10.1016/j.ijindorg.2010.03.005

Clerides, S. K., Lach, S., & Tybout, J. R. (1998). Is Learning by Exporting Important? Micro-Dynamic Evidence from Colombia, Mexico, and Morocco*. *The Quarterly Journal of Economics, 113*(3), 903-947. doi:10.1162/003355398555784

Dai, M., Huang, W., & Zhang, Y. (2020). Persistent effects of initial labor market conditions: The case of China's tariff liberalization after WTO accession. *Journal of Economic Behavior & Organization, 178*, 566-581. doi:https://doi.org/10.1016/j.jebo.2020.07.036

Dai, M., Huang, W., & Zhang, Y. (2021). How do households adjust to tariff liberalization? Evidence from China's WTO accession. *Journal of Development Economics, 150*, 102628. doi:https://doi.org/10.1016/j.jdeveco.2021.102628

Dai, M., Liu, H., & Lin, L. (2020). How innovation impacts firms' export survival: Does export mode matter? *The World Economy, 43*(1), 81-113. doi:https://doi.org/10.1111/twec.12847

Damijan, J. P., Kostevc, Č., & Polanec, S. (2010). From Innovation to Exporting or Vice Versa? *The World Economy, 33*(3), 374-398. doi:https://doi.org/10.1111/j.1467-9701.2010.01260.x

Das, S., Roberts, M. J., & Tybout, J. R. (2007). Market Entry Costs, Producer Heterogeneity, and Export Dynamics. *Econometrica, 75*(3), 837-873. doi:10.1111/j.1468-0262.2007.00769.x

De Fuentes, C., Niosi, J., & Peerally, J. A. (2021). Exploring innovation and export interplay in Canadian firms. *Economics of Innovation and New Technology, 30*(8), 786-806. doi:10.1080/10438599.2020.1786999

De Loecker, J. (2007). Do exports generate higher productivity? Evidence from Slovenia. *Journal of International Economics, 73*(1), 69-98. doi:https://doi.org/10.1016/j.jinteco.2007.03.003

De Loecker, J. (2013). Detecting Learning by Exporting. *American Economic Journal: Microeconomics, 5*(3), 1-21. doi:10.1257/mic.5.3.1

Ederington, J., & McCalman, P. (2008). Endogenous firm heterogeneity and the dynamics of trade liberalization. *Journal of International Economics, 74*(2), 422-440. doi:https://doi.org/10.1016/j.jinteco.2007.07.001

Egger, P., & Keuschnigg, C. (2015). Innovation, Trade, and Finance. *American Economic Journal: Microeconomics, 7*(2), 121-157. doi:doi: 10.1257/mic.20120032

Feng, L., Li, Z., & Swenson, D. L. (2017). Trade policy uncertainty and exports: Evidence from China's WTO accession. *Journal of International Economics, 106*, 20-36. doi:https://doi.org/10.1016/j.jinteco.2016.12.009

Heckman, J. (1979). Sample Selection Bias as a Specification Error. *Econometrica, 47*(1), 153-161. doi:10.2307/1912352

Hotz, V. J., & Miller, R. A. (1993). Conditional Choice Probabilities and the Estimation of Dynamic Models. *The Review of Economic Studies, 60*(3), 497-529. doi:10.2307/2298122

Imai, S., Jain, N., & Ching, A. (2009). Bayesian Estimation of Dynamic Discrete Choice Models. *Econometrica, 77*(6), 1865-1899. doi:https://doi.org/10.3982/ECTA5658

Impullitti, G., & Licandro, O. (2018). Trade, Firm Selection and Innovation: The Competition Channel. *The Economic Journal, 128*(608), 189-229. doi:https://doi.org/10.1111/ecoj.12466

Kalouptsidi, M., Scott, P. T., & Souza-Rodrigues, E. (2021). Identification of counterfactuals in dynamic discrete choice models. *Quantitative Economics, 12*(2), 351-403. doi:https://doi.org/10.3982/QE1253

Keane, M. P., & Sauer, R. M. (2009). Classification Error in Dynamic Discrete Choice Models: Implications for Female Labor Supply Behavior. *Econometrica, 77*(3), 975-991. doi:https://doi.org/10.3982/ECTA7642

Keane, M. P., & Wolpin, K. I. (2009). Empirical applications of discrete choice dynamic programming models. *Review of Economic Dynamics, 12*(1), 1-22. doi:https://doi.org/10.1016/j.red.2008.07.001

Lileeva, A., & Trefler, D. (2010). Improved Access to Foreign Markets Raises Plant-level Productivity for Some Plants. *The Quarterly Journal of Economics, 125*(3), 1051-1099. doi:10.1162/qjec.2010.125.3.1051

Lin, F., & Long, C. X. (2020). The impact of globalization on youth education: Empirical evidence from China's WTO accession. *Journal of Economic Behavior & Organization, 178*, 820-839. doi:https://doi.org/10.1016/j.jebo.2020.08.024

Lin, H.-l., & Lin, E. S. (2010). FDI, Trade, and Product Innovation: Theory and Evidence. *Southern Economic Journal, 77*(2), 434-464. doi:https://doi.org/10.4284/sej.2010.77.2.434

Liu, Q., & Ma, H. (2020). Trade policy uncertainty and innovation: Firm level evidence from China's WTO accession. *Journal of International Economics, 127*, 103387. doi:https://doi.org/10.1016/j.jinteco.2020.103387

Lu, Y., & Yu, L. (2015). Trade Liberalization and Markup Dispersion: Evidence from China's WTO Accession. *American Economic Journal: Applied Economics, 7*(4), 221-253. doi:10.1257/app.20140350

Melitz, M. J., & Redding, S. J. (2021). Trade and Innovation. *National Bureau of Economic Research Working Paper Series, No. 28945*. doi:10.3386/w28945

OECD, & Eurostat. (2005). *Oslo Manual: Guidelines for Collecting and Interpreting Innovation Data* (3rd Editio ed.).

Olabisi, M. (2017). The Impact of Exporting and Foreign Direct Investment on Product Innovation: Evidence from Chinese Manufacturers. *Contemporary Economic Policy, 35*(4), 735-750. doi:https://doi.org/10.1111/coep.12227

Olley, G. S., & Pakes, A. (1996). The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica, 64*(6), 1263-1297. doi:10.2307/2171831

Rust, J. (1987). Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher. *Econometrica, 55*(5), 999-1033. doi:10.2307/1911259

Salomon, R. M., & Shaver, J. M. (2005). Learning by Exporting: New Insights from Examining Firm Innovation. *Journal of Economics & Management Strategy, 14*(2), 431-460. doi:https://doi.org/10.1111/j.1530-9134.2005.00047.x

Sun, S. (2023). Firm heterogeneity, worker training and labor productivity: the role of endogenous self-selection. *Journal of Productivity Analysis, 59*(2), 121-133. doi:10.1007/s11123-022-00652-1

Sun, S., & Anwar, S. (2024). Can We Reliably Identify the CES Preference Parameter from Firm Revenue and Cost Data? Evidence from Monte Carlo Experiments. *Journal of Quantitative Economics*. doi:10.1007/s40953-024-00397-8

Xiang, Y., Gubian, S., Suomela, B., & Hoeng, J. (2013). Generalized Simulated Annealing for Global Optimization: The GenSA Package. *The R Journal, 5*(1), 13-28. doi:10.32614/RJ-2013-002