Productivity Gains from International Trade: Does Firm Age Matter?*

M. Jahangir Alam†
HEC Montréal

August 1, 2018

Abstract

It has been well established that international trade generates productivity gains within industries by reallocating inputs from low-productivity to high-productivity firms. However, the literature does not differentiate between *young economies*–that is, economies with a relatively large proportion of young firms–and *old economies*. To compare the productivity gains between young and old economies, I use European firm-level data during the period 2006 to 2014. The results of my study support the common finding in the extant literature that productivity gains are higher in any economy with more international trade. However, I also find that such gains are approximately five times higher for younger economies. I show that this is due to higher degrees of productivity dispersion within industries in younger economies.

Keywords: International trade, Productivity, Reallocation, Firm age, Young economies

JEL classification: F14, O47, O52

*I am grateful to Joanne Roberts and Trevor Tombe for their valuable comments. I would like to thank Eugene Beaulieu, Curtis Eaton, Benoit Dostie, Wulong Gu, and Lars Vilhuber for their helpful comments and discussions. I would also like to thank seminar participants at the University of Calgary for their helpful inputs. This research was made possible through the use of Cornell University's Economics Compute Cluster Organization, which was partially funded through NSF Grant #0922005. Last but not least, I would like to thank Bureau van Dijk for access to the Amadeus database.

†Department of Applied Economics, HEC Montréal, Canada (e-mail: jahangir.alam@hec.ca).

1 Introduction

It has been well established that international trade generates productivity gains at the industry level by shifting resources away from low-productivity firms and toward high-productivity firms (Pavcnik, 2002; Melitz, 2003; Trefler, 2004). However, the literature does not differentiate between *young economies*–that is, economies with a relatively large proportion of young firms–and *old economies* while estimating productivity gains from international trade. Firm age may matter for productivity gains from international trade for two reasons: 1) young firms are more likely to fail, and 2) the young firms that do survive can grow quickly (Bartelsman et al., 2009; Haltiwanger et al., 2013), because it is easier and cheaper for young firms to reallocate inputs. In other words, their reallocation channel may be stronger, and their productivity gains from international trade may be higher. Hence the question: Does increased exposure to foreign competition generate more productivity gains from between-firm reallocation in relatively younger economies?

Using European firm-level data, I find that productivity gains, from between-firm reallocation within industries, from international trade for young economies are indeed higher than for old economies.[1] Specifically, productivity gains from international trade for young economies are approximately five times higher than for old economies. To explain the reason for this result, I show that industries in young economies exhibit higher degrees of productivity dispersion.

To establish my results, I estimate the effect of the interaction between mean firm age and trade share on productivity gains from between-firm reallocation within industries. However, the effect of the interaction on the productivity gains may not be identified, because the mean firm age could be endogenous: unobserved productivity shocks could be affecting both, the mean firm age and the productivity gains. For example, by promoting competition and creating high exit and entry of firms within industries, some regulatory policies may affect mean firm age. Moreover, these policies may also directly affect productivity gains by promoting the reallocation of inputs from less to more productive firms. To address the

[1] From here on, the term "productivity gains" in this paper, refers to percentage changes in the aggregate productivity from reallocating inputs within industries for a given country, relative to 2006. In this study, I consider two main productivity measures: revenue labor productivity and total factor productivity (TFP). Since my result holds for both productivity measures, I use them interchangeably.

possible endogeneity of mean firm age, I use the instrumental variable (IV) method. As an instrument for mean firm age, I use the number of years after the birth of the current form of government. Since a change in the form of government can bring institutional changes (Congleton and Yoo, 2014; Fors and Olsson, 2007) that directly affect productivity gains from reallocating inputs, I add a vector for the institutional quality, to the estimation.

I use the Amadeus database, which covers European countries, during the period 2006 to 2014. This database is uniquely well-suited for this research problem because it has firm-level data for both young and old economies. Due to high exit and entry of firms after the collapse of Communism in East European countries, the average age of firms is lower in those countries than in West European countries. To compare productivity gains from reallocation inputs from international trade between young and old economies, I consider East European countries as young economies and West European countries as old economies.

To calculate productivity gains from between-firm reallocation within industries including the contribution of exiting and entering firms, I use the Dynamic Olley-Pakes Decomposition method that is developed by Melitz and Polanec (2015). I find that the contribution of between-firm reallocation to aggregate TFP growth is approximately 27 percent. My results are consistent with the growing body of literature that empirically estimates the reallocation contributions to aggregate productivity growth. For example, Pavcnik (2002) calculates that the aggregate productivity in Chile has increased by 19 percent during the period 1979 to 1986, of which approximately 12.7 percent was due to productivity gains from reallocating inputs between plants. In addition, Fernandes (2007), using data from Colombia during the period 1977 to 1991, provides evidence that, on average, 32 percent of change in industry productivity was due to TFP gains between plants. Using the results from Trefler (2004) and Lileeva and Trefler (2010), Melitz and Trefler (2012) report that between-plant contribution to total labor productivity gains from the Canada-U.S. free trade agreement is around 8.4 percent during the period 1988 to 1996. Using Slovenian manufacturing data from during the period 1995 to 2000, Melitz and Polanec (2015) find that the between-firms contribution to the aggregate TFP growth (considering only surviving firms) is approximately 23 percent.

My result supports the argument advanced by Melitz (2003) that productivity gains at the industry level from reallocating inputs are higher with more international trade. The magnitude of this result is comparable with other studies in the trade literature. For example, the magnitude of the international trade effect on productivity gains at the industry level from reallocating inputs that I estimate (0.40 percent per year) for European countries is lower, as expected, than the estimate of Jonsson and Subramanian (2001) (0.82 percent per year) for South African countries.

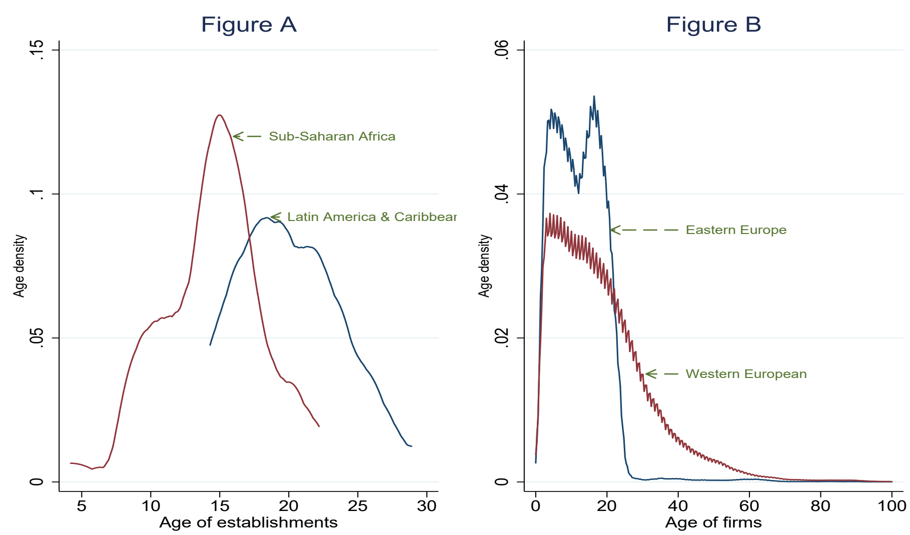
The rest of this paper is organized as follows. Section 2 provides two stylized facts for motivation. Section 3 provides a description of my methods, including the method for measuring productivity and international trade. Section 3 also provides descriptions of the decomposition method, and regression model used to test my hypothesis that productivity gains from international trade are higher for younger economies. Section 4 provides a description of the data and some descriptive statistics. Section 5 presents and discusses the results. Section 6 provides the results of three robustness experiments: measuring age of firms by median age of firms instead of mean age of firms, estimating revenue labor productivity as an alternative measure of productivity, and imputing missing observations. This paper concludes with Section 7, where I suggest future research possibilities.

2 Motivation

In this section, I provide two stylized facts for motivation. First, I show that the mean firm age differs across countries. Second, I provide evidence of the negative correlation between the mean firm age and variance of productivity.

To show that the mean firm age differs across countries, I use establishment-level data from the World Bank Enterprise Surveys for the period between 2006 and 2014. Figure 1 shows that the average age of establishments is lower in Sub-Saharan countries than in Latin American and Caribbean countries. Since my study focuses on European countries, I compare the age distribution of firms in East European countries with that of West European countries using the Amadeus database. Figure 1 also shows that the average age of firms is lower in East European counties when compared to West European countries. This difference could be due to the high exit and entry of firms after the collapse of Communism. Based on these two facts, I can conclude that the inference that there are younger firms in younger economies holds more generally.

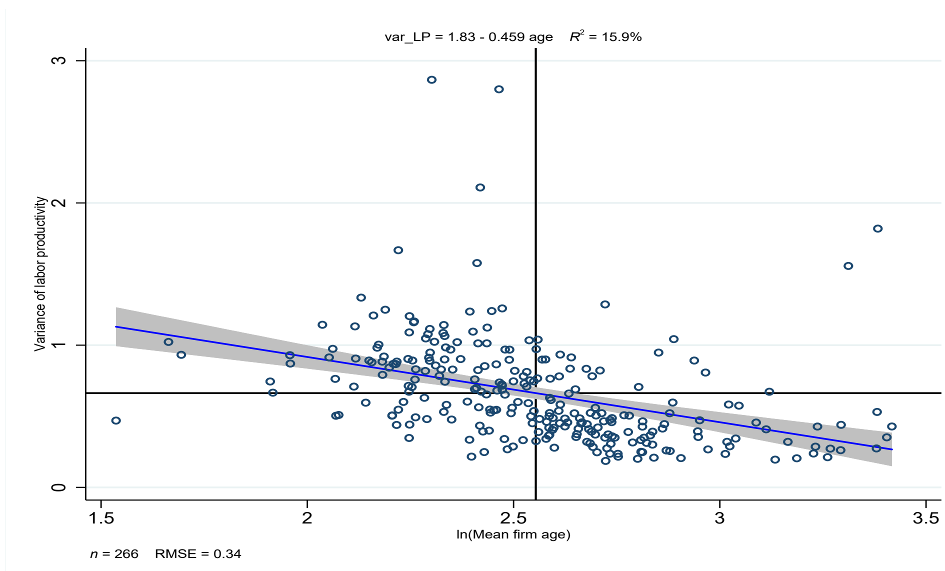
Figure 1: Age distribution of firms



Sources: Amadeus database (2006 - 2014) and World Bank Enterprise Surveys (2006-2014)

Note: Figure A is based on the World Bank Enterprise Surveys and Figure B is based on the Amadeus database. For figure B, I drop firms with the age of more than 100 years.

Figure 2: Correlation between variance of firm productivity and mean firm age



Source: Amadeus database (2006 - 2014)

Note: Each point represents various two-digit industries, and the shaded area shows the confidence interval of the fitted line. Both reference lines are set at the mean. The figure shows that countries with a large proportion of young firms have a high variance of firm productivity.

To provide evidence of the negative correlation between mean firm age and variance of productivity, I calculate the variance of firm productivity and mean firm age for various two-digit industries within a country and present it in Figure 2. This negative correlation[2] means that young economies (young industries) have a higher variance of firm productivity than old economies (old industries). The significance of this fact is that young economies have a higher scope for productivity gains from reallocating inputs.

Based on these two stylized facts, I hypothesize that the reallocation channel is stronger in young economies. As a result, I expect productivity gains from international trade to be higher in young economies.

3 Methods

In this section, I explain the method for measuring productivity and international trade. There are numerous possibilities for the choice of a productivity measure and an international trade measure. In addition, I describe the method used to decompose the aggregate productivity changes for calculating the productivity gains from reallocating inputs. I also explain the regression model to test my hypothesis that productivity gains from international trade are higher for relatively young economies, than for old economies.

3.1 Choice of Productivity Measure and Weight

There are numerous possibilities for the choice of a productivity measure and its associated market share weight. I consider two main productivity measures and their associated weights: one measure of revenue productivity with employment shares as weights, and another measure of TFP with value-added shares as weights. I directly compute the revenue productivity as the log of value added per unit of labor. For my second measure, I estimate TFP using the Wooldridge (2009) method.[3] I estimate TFP as the residual of the firm-level

[2] I also find the negative correlation between median firm age and variance of firm productivity.

[3] Olley and Pakes (1996) develop an estimator that uses investment as a proxy for unobservable productivity shocks. Levinsohn and Petrin (2003) propose a modification of the Olley and Pakes (1996) approach to address the problem of lumpy investment and suggest using intermediate inputs as a proxy for unobserved productivity. While Levinsohn and Petrin (2003) invert the intermediate input demand functions that are unconditional on the labor input, Ackerberg et al. (2015) argue that the moment condition underlying the

production function separately for all NACE 2-digit industries:

$$logv_{it} \quad = \quad \beta_l logl_{it} + \beta_k logk_{it} + log\omega_{it} + \epsilon_{it} \tag{1}$$

where v_{it}, l_{it}, and k_{it} denote the real value added, number of employees, and real capital of firm i in period t, and β_l and β_k denote the coefficients for labor and capital.

3.2 Choice of International Trade Measure

To measure international trade, I use trade openness measured by trade share, which is calculated by adding exports to imports and dividing the sum by output. I use imports, exports, and output data from the World Input Output Tables (WIOT) at a two-digit industry-level.

3.3 Choice of Decomposition Method

There are numerous possibilities when it comes to the choice of method for the decomposition of the aggregate productivity changes. I use the Dynamic Olley-Pakes Decomposition method[4] because it measures the contribution of entering and exiting firms to aggregate industry productivity changes properly. This method was developed by Melitz and Polanec (2015), based on the decomposition method of Olley and Pakes (1996).

The Dynamic Olley-Pakes Decomposition method has four components: productivity shifts among surviving firms, market share reallocations among surviving firms, the contribution of entering firms, and the contribution of exiting firms. To calculate productivity gains from reallocating inputs among firms, I first explain the Dynamic Olley-Pakes Decomposition method. In explaining this method, I define industry productivity as a weighted

first stage estimating equation does not identify the labor coefficient and suggest inverting intermediate input demand functions that are conditional on the labor input. Wooldridge (2009) proposes estimating the first and second stage moments in Levinsohn and Petrin (2003) structure simultaneously.

[4]I use the Dynamic Olley-Pakes Decomposition method rather than the Griliches and Regev (1995) method or the Foster et al. (2001) method, as both are based on the Baily et al. (1992) decomposition method. Melitz and Polanec (2015) argue that these two decomposition methods suffer from some biases that stem from their construction method. In general, the theoretical direction of these biases (as well as their magnitudes) is ambiguous. Theoretically, Melitz and Polanec (2015) show that these biases involve an over-measurement of the contribution of entry. In addition, Melitz and Polanec (2015) show empirically that the magnitude of these biases are substantial and are then also reflected in a substantial under-measurement of the contribution of surviving firms to aggregate productivity growth.

average of firm-level productivity:

$$\Phi_t = \sum_{i=1}^{n_t} s_{it}\varphi_{it} \tag{2}$$

where Φ_t is the industry productivity at time t, φ_{it} is the productivity of firm i at time t, and s_{it} is the market share of firm i at time t ($s_{it} \geq 0$ and $\sum_{i=1}^{n_t} s_{it} = 1$). Then, I decompose the industry productivity into two terms as follows:

$$\Phi_t = \sum_{i=1}^{n_t}(\bar{s}_t + \Delta s_{it})(\bar{\varphi}_t + \Delta\varphi_{it}) \tag{3}$$

where $\Delta s_{it} = s_{it} - \bar{s}_t$, $\Delta\varphi_{it} = \varphi_{it} - \bar{\varphi}_t$, $\bar{\varphi}_t$ is the unweighted mean productivity of firms at time t ($=\frac{1}{n_t}\sum_{i=1}^{n_t}\varphi_{it}$) and $\bar{s}_t$ is the mean market share at time t. From the equation (3), I can derive the following equations:

$$\begin{aligned} \Phi_t &= n_t\bar{s}_t\bar{\varphi}_t + \sum_{i=1}^{n_t}\Delta s_{it}\Delta\varphi_{it} \\ &= \bar{\varphi}_t + \sum_{i=1}^{n_t}(s_{it} - \bar{s}_t)(\varphi_{it} - \bar{\varphi}_t) \\ &= \bar{\varphi}_t + cov(s_{it}, \varphi_{it}) \end{aligned} \tag{4}$$

In the equation (4), the aggregate industry productivity at time t is equal to the unweighted mean productivity of firms at time t plus the covariance between the market share and firm productivity. Positive covariance between the market share and firm productivity means that factors of production are reallocated from less to more efficient firms.

To calculate the aggregate industry productivity change, I consider two periods: period t ($t = 2007$ to 2014) and the base year, period 0 (2006). With these two periods, I can divide firms into three groups. Group 1 consists of surviving firms that exist in both periods. Group 2 includes exiting firms that survive in period 0, but not in period t. Group 3 comprises entering firms that exist in period t, but not in period 0. Then, the aggregate industry productivity change over period 0 and t is as follows:

$$\Delta\Phi_t = \Phi_t - \Phi_0 \tag{5}$$

where $\Delta\Phi_t$ is the aggregate industry productivity change, Φ_0 is the aggregate industry productivity in period 0, and Φ_t is the aggregate industry productivity in period t.

In period 0, the aggregate industry productivity of surviving and exiting firms is as follows:

$$\begin{aligned}
\Phi_0 &= s_{S_0}\Phi_{S_0} + s_{X_0}\Phi_{X_0}, \qquad [s_{S_0} + s_{X_0} = 1] \\
&= \Phi_{S_0} + s_{X_0}(\Phi_{X_0} - \Phi_{S_0}) \\
&= [\bar{\varphi}_{S_0} + cov_{S_0}(s_{i0}, \varphi_{i0})] + s_{X_0}[\bar{\varphi}_{X_0} + cov_{X_0}(s_{i0}, \varphi_{i0}) - \bar{\varphi}_{S_0} - cov_{S_0}(s_{i0}, \varphi_{i0})] \qquad (6)
\end{aligned}$$

In period t, the aggregate industry productivity of surviving and entering firms is:

$$\begin{aligned}
\Phi_t &= s_{S_t}\Phi_{S_t} + s_{E_t}\Phi_{E_t}, \qquad [s_{S_t} + s_{E_t} = 1] \\
&= \Phi_{S_t} + s_{E_t}(\Phi_{E_t} - \Phi_{S_t}) \\
&= [\bar{\varphi}_{S_t} + cov_{S_t}(s_{it}, \varphi_{it})] + s_{E_t}[\bar{\varphi}_{E_t} + cov_{E_t}(s_{it}, \varphi_{it}) - \bar{\varphi}_{S_t} - cov_{S_t}(s_{it}, \varphi_{it})] \qquad (7)
\end{aligned}$$

where Φ_{S_0} and Φ_{S_t} are the aggregate industry productivity of surviving firms at time 0 and t respectively, Φ_{X_0} is the aggregate industry productivity of exiting firms at time 0, and Φ_{E_t} is the aggregate industry productivity of entering firms at time t. In addition, s_{S_0} and s_{S_t} are the aggregate market share of surviving firms at time 0 and t respectively, s_{X_0} is the aggregate market share of exiting firms at time 0, and s_{E_t} is the aggregate market share of entering firms at time t. Moreover, $\bar{\varphi}_{S_0}$ and $\bar{\varphi}_{S_t}$ are the unweighted mean productivity of surviving firms at time 0 and t respectively, $\bar{\varphi}_{X_0}$ is the unweighted mean productivity of exiting firms at time 0, and $\bar{\varphi}_{E_t}$ is the unweighted mean productivity of entering firms at time t. Furthermore, cov_{S_0} and cov_{S_t} are the covariance of surviving firms at time 0 and t respectively, cov_{X_0} is the covariance of exiting firms at time 0, and cov_{E_t} is the covariance of entering firms at time t.

Then, I substitute the equations (6) and (7) into equation (5):

$$\begin{aligned}
\Delta\Phi_t = \Phi_t - \Phi_0 &= (\bar{\varphi}_{S_t} - \bar{\varphi}_{S_0}) + (cov_{S_t} - cov_{S_0}) \\
&+ s_{E_t}[(\bar{\varphi}_{E_t} - \bar{\varphi}_{S_t}) + (cov_{E_t} - cov_{S_t})] \\
&+ s_{X_0}[(\bar{\varphi}_{S_0} - \bar{\varphi}_{X_0}) + (cov_{S_0} - cov_{X_0})] \qquad (8)
\end{aligned}$$

where $\bar{\varphi}_{S_t} - \bar{\varphi}_{S_0}$ represents the productivity distribution shifts among surviving firms. This is referred to as the change of unweighted mean productivity, which is the productivity change within-firms. $cov_{S_t} - cov_{S_0}$ represents the market share reallocation contribution (covariance change) of surviving firms to the aggregate industry productivity change. $s_{E_t}[(\bar{\varphi}_{E_t} - \bar{\varphi}_{S_t}) + (cov_{E_t} - cov_{S_t})]$ measures the reallocation contribution of entering firms to the aggregate industry productivity change and $s_{X_0}[(\bar{\varphi}_{S_0} - \bar{\varphi}_{X_0}) + (cov_{S_0} - cov_{X_0})]$ represents the reallocation contribution of exiting firms to the aggregate industry productivity change.

In the equation (8), the aggregate industry productivity changes decomposes into three components for three groups of firms: surviving, exiting, and entering firms. The first line represents the contribution of surviving firms to the aggregate industry productivity change. The second line represents the contribution of entering firms. The third line represents the contribution of exiting firms. There are two terms for those surviving firms: one for the productivity change within-firm and another for the reallocation contribution.

3.4 Productivity Gains from Reallocating Inputs

Unlike previous studies, to measure the productivity gains from reallocating inputs within industries for a given country using the equation 8, I add the productivity contribution of both exiting and entering firms to the contribution of market share reallocations among surviving firms:

$$\begin{aligned} ln\left(\frac{\Omega_{ict}}{\Omega_{ic0}}\right) &= (cov_{S_t} - cov_{S_0}) + s_{E_t}[(\bar{\varphi}_{E_t} - \bar{\varphi}_{S_t}) + (cov_{E_t} - cov_{S_t})] \\ &+ s_{X_0}[(\bar{\varphi}_{S_0} - \bar{\varphi}_{X_0}) + (cov_{S_0} - cov_{X_0})] \end{aligned} \tag{9}$$

where $ln(\frac{\Omega_{ict}}{\Omega_{ic0}})$ represents the productivity gains from reallocating inputs within industries; Ω_{ict} represents the productivity gains from reallocating inputs for period t (2007 to 2014); and Ω_{ic0} represents the productivity gains from reallocating inputs for the base year, period 0 (2006). In my estimation, I use the calculated value of equation (9) as a dependent variable.

3.5 Regression Model

To provide evidence that productivity gains from international trade are higher for young economies than for old economies, I estimate the effect of the interaction between mean firm age and trade share on the productivity gains from reallocating inputs within industries, using the following regression:

$$ln\left(\frac{\Omega_{ict}}{\Omega_{ic0}}\right) = \gamma + \theta A_{ict} + \lambda T_{ict} + \delta T_{ict} * A_{ict} + X_{it}^{T}\beta + \alpha_i + \alpha_t + \alpha_c + \upsilon_{ict} \tag{10}$$

where $\left(\frac{\Omega_{ict}}{\Omega_{ic0}}\right)$ is the productivity gain from reallocating inputs of industry i for country c at time t relative to the base period, A_{ict} is the mean firm age of industry i for country c at time t, T_{ict} is the trade openness (measured by trade share as mentioned before) of industry i for country c at time t, X_{it} refers to the other controls (Herfindahl index and institutional quality) of industry i for country c at time t, α_i is industry fixed effects, α_t is time fixed effects, α_c is country fixed effects, and υ_{ict} is the error term.

As mentioned before, I measure the trade share by adding exports to imports and dividing the sum by output. I include the normalized Herfindahl index[5] (an indicator of the level of competition among firms) of the industry to capture scale effects, industry dummies to capture the effects of unobservable and time-invariant industry characteristics, year dummies to capture time trends, and country dummies to capture country-specific fixed effects. The parameter of interest is δ, which explains how countries with a larger proportion of young firms, and who are more open to trade, could have more productivity gains. However, this specification may not identify the effect of the interaction (δ) on the productivity gains from reallocating inputs. The problem is that mean firm age may be endogenous: unobserved productivity shocks could be affecting both the mean firm age and productivity gains from reallocating inputs. For example, by promoting competition and creating a high exit and entry of firms within industries, some regulatory policies may affect mean firm age. Moreover, these policies may also directly affect productivity gains by promoting the reallocation of

[5]To control for market concentration, I use the normalized Herfindahl index, $H^* = \frac{H-\frac{1}{N}}{1-\frac{1}{N}}$, where $H = \sum_{j=1}^{N} S_j^2$, N is the number of plants within industries, and S_j is the market share measured by value added of plant j within industries.

inputs from less to more productive firms.

To overcome the problem posed by the possible endogeneity of mean firm age, I use the instrumental variable (IV) method. I use the number of years after the birth of the current form of government as the instrument for mean firm age. Since a change in the form of government can bring institutional changes that affect productivity gains from reallocating inputs, I add a vector for the institutional quality to this estimation. These indicators are voice and accountability, political stability and absence of violence, government effectiveness, the regulatory quality, rule of law, and control of corruption. Moreover, to create an instrument for the interaction term, I multiply the instrument for mean firm age by trade openness. My IV estimates could have two issues. First, industries with a large proportion of young firms are more likely to trade less than industries with a large proportion of old firms. Second, there is a potential endogeneity problem between productivity gains from reallocating inputs and international trade because of reverse causality.

4 Data and Descriptive Statistics

In this section, I present a short description of my data sources and descriptive statistics to provide insights into my results.

4.1 Sources of Data

In this paper, I employ various sources of data, which are listed in Table 1. As a primary database, I use firm-level data from the Amadeus database.[6] I use this database through the Wharton Research Data Services (WRDS).[7] The Amadeus database is compiled by the Bureau van Dijk (BvD), and covers firms in Europe reporting to either the local tax authorities or data collection agencies. This database contains financial information on millions of public and private firms in European countries. One advantage of focusing on European countries is that I can observe both young countries (i.e., East European countries) and old countries (i.e., West European countries). I use 14 European countries (Table 2), which have

[6] `https://amadeus.bvdinfo.com/version-2015813/home.serv?product=amadeusneo`
[7] `https://wrds-web.wharton.upenn.edu/wrds/`

sufficient observations to implement the Dynamic Olley-Pakes Decomposition method. Since the number of observations for exiting and entering firms before 2006 and after 2014 is low, I use the period of 2006 to 2014 in order to cover the maximum number of observations in each year.

Table 1: The sources of data

Data	**Data Source**	**Data Level**
Primary database	Amadeus database	Firm-level
Exports, Imports, and Output	World Input Output Table (WIOT) database	NACE2
Deflators	Structural Analysis (STAN) database	SIC2
Nominal exchange rate	Penn World Table 9.0	Country-level
TFP	Penn World Table 9.0	Country-level
Governance Indicators	World Bank Governance Indicators database	Country-level
Year of the current form of government	Wikipedia website	Country-level
Age of establishment	World Bank Enterprise Surveys	Establishment-level

From the Amadeus database, I use sales, the number of employees, tangible fixed assets, material costs, NACE Rev. 2, the year of closing, the year of incorporation, the size of firms, type of legal form of firms, and legal status of firms. To calculate the capital stock, I use tangible fixed assets. I construct the age of firms using the standard measure: the difference between the year of the balance sheet information and the year of the incorporation of the firm plus one. Finally, I calculate the value added as the difference between operating revenue and material costs.

To compute the trade share at a two-digit industry-level, I use the National Input Output Tables from the World Input Output Table (WIOT) (Timmer et al., 2015).[8] From the National Input Output Tables, I use exports, imports, and output at a two-digit industry-level. I drop the coke and refined petroleum products industry (NACE 19) for Slovenia because the average amount of exports plus imports, for my study period, is 241 times higher than the output.

To deflate firm value added, tangible fixed assets, and material costs, I use deflators from

[8] http://www.wiod.org/new_site/home.htm

the STAN database from the Organization for Economic Co-operation and Development (OECD).[9] To convert all monetary values into the same currency, I use the exchange rate from the Penn World Table 9.0,(Feenstra et al., 2015),[10]. To check the sample representativeness of my estimated TFP, I also use the Penn World Table (PWT) 9.0. To measure the institutional quality, I use the World Bank Governance Indicators (Kaufmann and Kraay (2018)).[11] These indicators are voice and accountability, political stability and absence of violence, government effectiveness, regulatory quality, rule of law, and control of corruption. For the instrumental variable, I use Wikipedia to collect the established year of the current form of government (Table 2). Finally, to show that mean firm age differs across countries, I use establishment-level data from the World Bank Enterprise Surveys for the period between 2006 and 2014.[12]

Table 2: Established year of the current form of government

West European Countries	**Established Year**	**East European Countries**	**Established Year**
France	1958	Bulgaria	1989
Germany	1949	Czech Republic	1993
Italy	1948	Estonia	1991
Portugal	1976	Hungary	1989
Spain	1978	Poland	1989
Sweden	1974	Romania	1989
		Slovakia	1993
		Slovenia	1991

Source: `https://en.wikipedia.org/wiki/List_of_sovereign_states_by_date_of_formation`, date: 23 February, 2017.
Note: For Estonia, I use the recent date of acquisition of sovereignty. For France, I use the year when the French Fifth Republic was formed.

4.2 Cleaning the Amadeus Database

I use only the unconsolidated account of firms for several reasons, the most important being to avoid including total sales of a multinational across affiliates located in different countries.

[9] For countries that are not part of OECD, I use the calculated mean deflators of European OECD countries.
[10] DOI: 10.15141/S5J01T
[11] `http://info.worldbank.org/governance/WGI/#home`
[12] Enterprise Surveys (`http://www.enterprisesurveys.org`), The World Bank

My analysis focuses on manufacturing industries since challenges related to estimating the production function are less severe in the manufacturing sector than in other sectors.

I clean the data in five steps. First, I clean the data of basic reporting mistakes. In this step, I drop firms if the number of employees is negative in any year, or if operating revenues are negative in any year, or if sales are negative in any year, or if total assets are negative in any year, or if tangible fixed assets are negative in any year, or if intangible fixed assets are negative in any year, or if age of firms is negative or missing in any year. In addition, I drop firm-year observations with zero values for operating revenues, the number of employees, total assets, and tangible fixed assets. Furthermore, I drop firm-year observations with missing or zero values for the costs of employees. I also drop firm-year observations with zero or negative values for material costs. Second, in order to include the same set of firms, I keep only those firms that generate positive value added. Third, I trim observations that are below the one percentile or above the 99 percentile of the distribution for value added, labor, tangible fixed assets, and material costs. Fourth, I drop firm-year observations with zero values for value added, number of employees, tangible fixed assets, and material costs. Table A1 shows the percentage of missing observations in value added, labor or number of employees, capital, and material costs.[13] Finally, I drop the industry and year pair if the number of observations for either survival group and exiting or survival group and entry group is five or less.

4.3 Descriptive Statistics

Table 3 and 4 report the summary statistics for the first and last years of my sample (2006 and 2014). Using the exchange rate from the PWT 9.0, I convert all monetary values into the same currency. Using deflators from the STAN database of the OECD, I deflate value added by the gross value added deflator at NACE two-digit, tangible fixed assets by the gross fixed capital formation deflator, and material costs by the gross intermediate input deflator at 2005 prices.

Table 3 depicts the number of observations for three groups of firms (i.e., surviving, exiting, and entering) in 2006 and 2014. The number of survival firms for each country in

[13] For robustness check, I impute these four variables.

2006 and 2014 are not the same because I create the indicator variable for those groups before cleaning and dropping observations. For example, after cleaning the number of survival firms for Bulgaria in 2006, it is 2,937 whereas in 2014, it is 7,144. Creating this indicator variable for those groups before dropping or cleaning would remove false death and birth from my analysis.

Table 3: Number of firms in 2006 and 2014

Country	**Year 2006**			**Year 2014**		
	Survival firms	Exiting firms	Total firms	Survival firms	Entering firms	Total firms
France	16951	11506	28457	14371	11187	25558
Germany	1084	1962	3046	1054	775	1829
Italy	35185	4611	39796	55044	37492	92536
Portugal	16855	3360	20215	15600	8706	24306
Spain	3077	986	4063	33495	17251	50746
Sweden	11191	435	11626	9610	4182	13792
Bulgaria	2937	235	3172	7144	6117	13261
Czech Republic	7380	1067	8447	5318	4349	9667
Estonia	1777	124	1901	1679	1409	3088
Hungary	340	59	399	1525	861	2386
Poland	2892	1542	4434	201	142	343
Romania	15024	3251	18275	13040	11681	24721
Slovakia	2403	590	2993	2288	3997	6285
Slovenia	2080	194	2274	2044	4195	6239

Note: Exiting, entering, and survival columns represent the number of observations.

As the Table 4 shows, Germany has the highest average firm age in both years whereas Slovakia has the lowest mean firm age in 2006 and Bulgaria has the lowest median firm age in 2014. Moreover, the average revenue productivity for Germany in both 2006 and 2014 is higher than that of the other countries. Furthermore, Sweden has the lowest variance of productivity in both 2006 and 2014 whereas Slovakia has the highest variance of labor productivity in 2006 and Bulgaria has the highest variance of labor productivity in 2014. These results imply that the average age of the firm is higher in West European countries and the variance of labor productivity is lower in those countries than in East European countries.

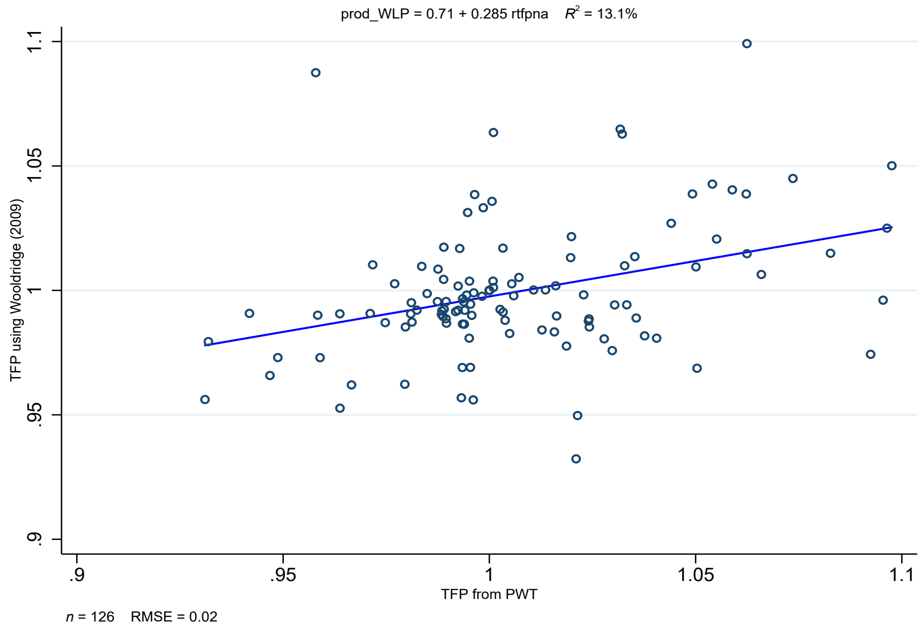
Table 4: Descriptive statistics in 2006 and 2014

Country	Year 2006					Year 2014				
	Labor	Age	Median Age	LP	Var(LP)	Labor	Age	Median Age	LP	Var(LP)
France	30.39	18.32	15	11.60	0.30	30.44	20.90	18	11.82	0.32
Germany	189.22	35.68	21	11.93	0.39	302.70	37.72	25	12.12	0.32
Italy	27.25	19.97	18	11.85	0.40	17.02	19.42	16	11.62	0.50
Portugal	17.44	15.65	13	10.31	0.43	15.91	17.72	14	10.37	0.55
Spain	89.72	24.82	22	11.70	0.28	14.41	19.40	19	11.14	0.43
Sweden	11.01	19.50	17	11.52	0.24	11.99	21.74	20	11.60	0.28
Bulgaria	82.35	16.42	11	9.36	1.09	28.17	11.31	7	9.20	0.95
Czech Republic	74.39	10.63	11	10.27	0.61	69.04	15.07	16	10.34	0.65
Estonia	25.40	9.61	9	9.82	0.58	15.53	11.91	10	10.19	0.56
Hungary	140.84	11.68	12	10.56	0.80	94.08	16.47	18	10.63	0.75
Poland	113.41	17.07	13	10.10	0.64	165.06	20.71	18	10.62	0.54
Romania	32.29	9.25	9	9.08	0.82	29.45	13.21	12	9.37	0.87
Slovakia	77.26	9.10	9	10.29	1.28	41.91	12.38	12	10.48	0.71
Slovenia	55.93	13.52	15	10.95	0.37	18.12	17.30	20	10.93	0.43

Note: Both labor productivity (LP) and variance of labor productivity (Var(LP)) are calculated within-industry. To aggregate labor productivity and variance of labor productivity, I use the employment share weights.

To check sample representativeness, I compare my estimated TFP with the aggregate TFP at constant national prices (2011=1) from the PWT 9.0. I convert my estimated TFP relative to 2011 to make data that is comparable with data from PWT. Figure 3 shows a positive correlation between my aggregate TFP and TFP from PWT (0.12). This positive correlation implies that my aggregate TFP estimates using firm-level data represents the aggregate TFP of selected countries for this study.

Figure 3: Correlation between estimated TFP and aggregate TFP from PWT



Source: TFP from PWT 9.0
Note: This graph compares estimated TFP based on the Wooldridge (2009) method with aggregate TFP at constant national prices (2011=1) from PWT 9.0. To make both TFPs comparable, I convert my estimated TFP relative to 2011.

Table 5 shows the contribution of within-firm growth and between-firm reallocation to aggregate labor productivity and TFP growth, which are calculated as relative shares to the total contribution of firms. This result is based on the Dynamic Olley-Pakes Decomposition method. The table shows that the contribution of between-firm reallocation to the aggregate TFP growth is around 27 percent during the period from 2007 to 2014 relative to 2006.

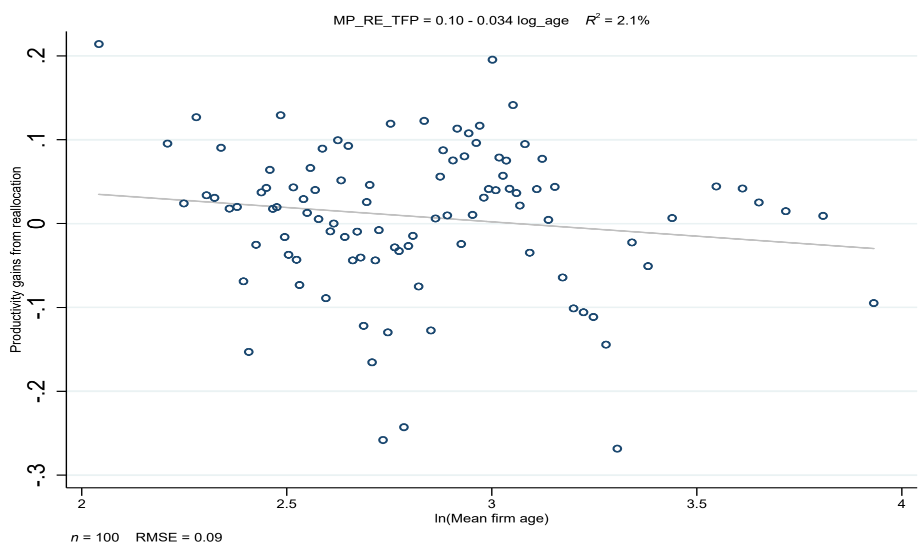
To show the correlation between the contribution of between-firm reallocation with mean firm age, I plot these two variables in Figure 4 and calculate the correlation in Appendix A2. Figure 4 exhibits a negative relationship between the productivity gains from reallocating inputs and mean firm age and Appendix A2 shows that the correlation between these two variables is significant. This negative relationship between these two variables means that young economies could have higher TFP gains by reallocating inputs from low to high productive firms. Due to this negative association, I expect that more international trade, which promotes the reallocation of inputs, leads to more TFP gains.

Table 5: Contributions of within-firm growth and between-firm reallocation (%)

Year	**Labor Productivity–Emp. Share Weights**		**TFP–Value-added Share Weights**	
	Within-Firm Growth	Between-Firm Reallocation	Within-Firm Growth	Between-Firm Reallocation
2007	92.02	7.98	101.74	-1.74
2008	86.55	13.45	86.55	13.45
2009	60.08	39.92	123.04	-23.04
2010	76.33	23.67	169.86	-69.86
2011	68.38	31.62	75.13	24.87
2012	35.05	64.95	192.30	-92.30
2013	55.32	44.68	32.47	67.53
2014	54.12	45.88	72.74	27.26

Note: The contributions of within-firm growth and between-firm reallocation are calculated as relative shares to the total contribution of firms.

Figure 4: TFP gains from reallocation and mean firm age



Note: I divide the mean firm age into 100 groups, and then calculate the mean of both TFP gains from reallocating inputs and mean firm age for those groups. This figure of 100 points shows that countries with a large proportion of young firms have higher productivity gains from reallocating inputs.

5 Results and Discussion

In this section, I establish the causal effect of trade openness on TFP gains from reallocating inputs within industries. In addition, I show that this causal effect of trade openness is relatively higher for young economies than for old economies. Specifically, I estimate the interaction effect of firm age and trade openness on TFP gains from reallocating inputs within industries. To show the direction of bias of OLS estimates for that causal effect, I present both the results of OLS and IV estimations.

Table 6 shows the causal effect of trade openness on TFP gains from reallocating inputs using both OLS and IV estimations.[14] The first two columns show OLS results and the last two columns show IV results. Since I assume that the instrument (the number of years after the birth of the current form of government) indirectly affects productivity gains through institutional changes, I include a vector for the institutional quality in both OLS and IV estimations. The results show that the coefficients do not significantly differ with and without including the institutional quality (by comparing column (1) with column (2)) in the estimation. As mentioned before, the estimate of interaction in the OLS estimation could be biased due to a missing variable. To overcome the problem posed by the possible endogeneity of mean firm age, I use the IV method. Due to the upward bias in the OLS estimation, the age and interaction coefficients of IV are lower with higher standard errors in all specifications (comparing a similar column for OLS estimates with the similar column for IV estimates).[15] Except the coefficient of age, all coefficients have similar directions or signs for both OLS and IV estimations. However, the magnitude of coefficients differs in both OLS and IV estimations.

My parameter of interest is the coefficient of interaction between age and trade openness, which is negative in all specifications. This negative interaction coefficient means that countries with industries that have a large proportion of young firms could have greater productivity gains from reallocating inputs than countries with industries that have a large proportion of old firms, if both those economies increase their international trade.

[14] Appendix A3 shows the estimates of the first-stage regression for IV estimations.

[15] To show the size of standard errors, I include robust standard errors and cluster standard errors at country and industry levels in Appendix A4.

Table 6: The effect of trade openness on TFP gains

Dependent variable: TFP gains from between-firm reallocation	**OLS Regressions**		**IV Regressions**	
	(1)	(2)	(1)	(2)
Age	0.26***	0.28***	-0.79**	-0.91**
	(0.05)	(0.05)	(0.38)	(0.36)
Openness	0.32***	0.32***	0.23**	0.22**
	(0.06)	(0.06)	(0.09)	(0.09)
Age*Openness	-0.10***	-0.10***	-0.07*	-0.07*
	(0.02)	(0.02)	(0.04)	(0.04)
Herfindahl index	-0.95***	-0.96***	-1.16***	-1.18***
	(0.10)	(0.10)	(0.14)	(0.14)
Institutional quality	No	Yes	No	Yes
Year	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes
Country	Yes	Yes	Yes	Yes
N	1973	1973	1973	1973
R squared	.31	.32		
Cragg-Donald F-stat			16.42	19.32

Notes: An observation is a country, a year, and an industry. The dependent variable, which is the percentage change in aggregate TFP from reallocating inputs within-industry relative to 2006, is calculated using the Dynamic Olley-Pakes Decomposition method. The age is in log scale. Openness is measured by trade share, which is calculated by adding exports to imports and dividing the sum by outputs. See the method section for detailed explanations of the Herfindahl index and institutional quality. Standard errors are in parentheses. Cragg-Donald F stat is a test for weak identification. ***, **, and * indicate statistically significant coefficients at 1%, 5%, and 10% percent levels, respectively.

To compare my results with those of other studies, I calculate the magnitude of the international trade effect on TFP gains from reallocating inputs. Even though the coefficients of OLS and IV estimates have similar directions for both, trade openness coefficient and interaction coefficient, my preferred specification is the IV estimation with the institutional quality (last column). Using the trade openness coefficient of 0.22, the interaction coefficient of -0.07, and a mean firm age coefficient of 2.82, I find that the causal effect of international trade on TFP gains from reallocating inputs is 3.22 percent (see in Table 7).[16] This positive value implies that more international trade leads to more productivity gains. This result supports the argument advanced by Melitz (2003). The magnitude of my results is comparable with those of other studies. For example, the magnitude of the international trade effect on TFP gains from reallocating inputs that I estimate (0.40 percent per year) for European countries is lower, as expected, than the estimate of Jonsson and Subramanian (2001) (0.82 percent per year) for South African countries.[17]

Table 7: The effect of trade openness on TFP gains (young versus old economies)

	Age of firms	Openness	Causal effect(%)	TFP gains(%)	Relative TFP gains
All countries	2.82	1.19	3.22	3.84	
Old economies	3.10	0.94	1.37	1.28	1.00
Young economies	2.61	1.39	4.63	6.44	5.02

Note: In my analysis, I consider East European countries as young economies and West European countries as old economies. Using the trade openness coefficient of 0.22, the interaction coefficient of -0.07, and a mean firm age coefficient of 2.82 for all countries, I calculate that causal effect of 3.22 percent. TFP gains are calculated by multiplying that causal effect by average trade openness of 1.19. Finally, I calculate relative TFP gains as TFP gains relative to TFP gains for old economies. The causal effect and TFP gains are in percentage.

To make a precise statement on the extent to which productivity gains from reallocating inputs are higher in countries with a large proportion of young firms than old firms, I calculate those gains from both groups of countries. In order to compare the TFP gains from reallocating inputs due to international trade between young and old economies, I consider East European countries as young economies and West European countries as old economies and calculate the TFP gains by multiplying the causal effect by trade openness. Finally, I calculate relative TFP gains as TFP gains relative to TFP gains for old economies. I find that those TFP gains for young and old economies are 6.44 and 1.28 percent, respectively (Table 7). This implies that TFP gains from reallocating inputs due to international trade are 5.02 ($= \frac{6.44}{1.28}$) times higher for young economies than for old economies.

[16] This calculation does not exactly match because of dropping decimal points.

[17] To calculate their estimate of 0.82 (=0.2726*3.0), I use the estimate of between-firms contribution to the aggregate TFP growth of 27.26 percent (see Table 5) and Jonsson and Subramanian's estimate of 3.0 for the effect of international trade on TFP growth of the manufacturing sector.

6 Robustness

My primary results are consistent across three robustness checks. First, I measure the age of firms as the median age of firms instead of the mean age of firms. Second, I estimate the causal effect of international trade on revenue labor productivity instated of TFP gains from reallocating inputs. Finally, I impute missing observations.

6.1 Median Age of Firms

In order to check the consistency of my primary results, I estimate the causal effect of trade openness on TFP gains from reallocating inputs within industries by measuring the age of firms as the median age of firms instead of the mean age of firms. The results of this robustness check are shown in Table 8, which shows that the estimates are similar to the primary results, except that IV estimates are not statistically significant. However, the coefficient of the interaction term between the age of firms and trade openness is negative. Thus, this robustness also confirms that TFP gains from reallocating inputs from international trade are higher for young economies than for old economies.

6.2 Revenue labor productivity

In addition to establishing that TFP gains are higher for young economies when compared to old economies, I check whether revenue labor productivity gains from reallocating inputs from international trade are higher for those young economies. As TFP gains from reallocating inputs, I use the similar specifications to establish revenue labor productivity gains. Table 9 shows the causal effect of trade openness on revenue labor productivity gains from reallocating inputs. Again, the IV estimation with the vector for institutional quality is my preferred specification (last column). I find that the causal effect of international trade on productivity gains from reallocating inputs is positive. This result implies that international trade, in general, leads to more revenue labor productivity gains. However, in addition, the interaction coefficient is negative, which further confirms that revenue labor productivity gains from international trade are higher for young economies than for old economies, if both those economies increase their international trade.

Table 8: The effect of trade openness on TFP gains

Dependent variable: TFP gains from between-firm reallocation				
	OLS Regressions		**IV Regressions**	
	(1)	(2)	(1)	(2)
Age	0.21***	0.23***	-1.04*	-1.12**
	(0.04)	(0.04)	(0.60)	(0.54)
Openness	0.31***	0.32***	0.17	0.15
	(0.07)	(0.07)	(0.19)	(0.19)
Age*Openness	-0.10***	-0.10***	-0.04	-0.03
	(0.03)	(0.03)	(0.08)	(0.08)
Herfindahl index	-0.95***	-0.96***	-1.20***	-1.21***
	(0.10)	(0.10)	(0.16)	(0.16)
Institutional quality	No	Yes	No	Yes
Year	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes
Country	Yes	Yes	Yes	Yes
N	1973	1973	1973	1973
R squared	.31	.32		
Cragg-Donald F-stat			5.82	7.84

Notes: An observation is a country, a year, and an industry. The dependent variable, which is the percentage change in aggregate TFP from reallocating inputs within-industry relative to 2006, is calculated using the Dynamic Olley-Pakes Decomposition method. The age is in log scale. Openness is measured by trade share, which is calculated by adding exports to imports and dividing the sum by outputs. See the method section for detailed explanations of the Herfindahl index and institutional quality. Standard errors are in parentheses. Cragg-Donald F stat is a test for weak identification. ***, **, and * indicate statistically significant coefficients at 1%, 5%, and 10% percent levels, respectively.

Table 9: The effect of trade openness on productivity gains

Dependent variable: Productivity gains from between-firm reallocation				
	OLS Regressions		**IV Regressions**	
	(1)	(2)	(1)	(2)
Age	0.23***	0.24***	0.13	-0.01
	(0.04)	(0.04)	(0.26)	(0.25)
Openness	0.23***	0.23***	0.20***	0.19***
	(0.05)	(0.05)	(0.07)	(0.07)
Age*Openness	-0.08***	-0.08***	-0.07***	-0.06**
	(0.02)	(0.02)	(0.03)	(0.03)
Herfindahl index	-0.89***	-0.90***	-0.91***	-0.95***
	(0.08)	(0.08)	(0.10)	(0.10)
Institutional quality	No	Yes	No	Yes
Year	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes
Country	Yes	Yes	Yes	Yes
N	2068	2068	2068	2068
R squared	.25	.26		
Cragg-Donald F-stat			17.99	21

Notes: An observation is a country, a year, and an industry. The dependent variable, which is the percentage change in aggregate revenue productivity from reallocating inputs within-industry relative to 2006, is calculated using the Dynamic Olley-Pakes Decomposition method. The age is in log scale. Openness is measured by trade share, which is calculated by adding exports to imports and dividing the sum by outputs. See the method section for detailed explanations of the Herfindahl index and institutional quality. Standard errors are in parentheses. Cragg-Donald F stat is a test for weak identification. ***, **, and * indicate statistically significant coefficients at 1%, 5%, and 10% percent levels, respectively.

6.3 Multiple imputation

As I explained before, Table A1 shows the percentage of missing on value added, labor, capital, material costs, and age of firms. I dropped those observations for which value added, labor, capital, material costs, and age of firms are missing in the analysis in the previous section. In this section, I check whether my results could change by imputing those missing observations. Before imputing those missing observations, I test, within industries, the covariate-dependent missingness (CDM) assumption, which is an extension of the Little test (see Li (2013) for details).[18] Appendix Table A5 shows the percentage of industries that have passed the CDM test under significance level 0.05 by size of firms, legal form, legal status, and the year as covariates.

I use multiple imputation for missing observations. Since fitted values from regression or ratios do not have an error term, a single imputation based on regression or ratio methods does not reflect the full uncertainty created by missing observations. To give a full representation of the uncertainty, multiple imputation is appropriate. In addition, single imputation is applicable if missing observations are less than 5 percent. In my case, since more than 5 percent of observations for value added, labor, capital, and material costs are missing in the Amadeus database, I use multiple imputation.

To implement multiple imputation, I use the Multivariate Imputation by Chained Equations (MICE) algorithm in R to fit the classification and regression trees (CART) model (White et al., 2013; Burgette and Reiter, 2010). The CART approximates the conditional distribution of a single variable using multiple predictors. To implement the CART, I order the variables in terms of increasing percentages of missing values. For the first variable in this ordering with missing data, I fit the tree on all other variables. Suppose, to impute value added, I use several potential predictors such as labor, capital stock, material costs, age of firms, size of firms, legal form, legal status, and year. I then impute labor using those same predictors including value added. Similarly, I impute capital stock, material costs, and age of firms.

At each stage in the tree-building process, the goal is to use predictors to divide the data into homogeneous groups. The CART algorithm searches through all the observed values of the predictors. The process continues recursively (I consider 10 iterations to help move the trees away from the initial starting values) on each branch of the tree until the terminal nodes contain some minimum number of homogeneous observations, which is 5 in my study. To make the imputation more homogeneous, I run the imputation process

[18]Little test implements the χ^2 test of the missing completely at random (MCAR) for multivariate quantitative data proposed by Little (1988). The test statistic is asymptotically χ^2 distributed under the null hypothesis that there are no differences between the means of different missing-value patterns. Rejection occurs if the null hypothesis provides sufficient evidence to indicate that data are not MCAR.

separately for each two-digit industry within a country. I repeat this entire process 10 times to generate 10 completed datasets. Each imputed dataset is analyzed separately and results are averaged to calculate TFP or revenue labor productivity gains from reallocating inputs from international trade.[19] I also show that TFP or revenue labor productivity gains are higher for young economies than for older economies.

After imputing the missing observations, I estimate both OLS and IV regressions for TFP gains and revenue productivity gains (Table 10). For both OLS and IV regressions, column (1) presents estimates without including the institutional quality and column (2) presents estimates including the institutional quality. The number of observations for TFP gains does not match with previous estimates because observations are dropped due to either the labor or capital coefficient being negative. The number of observations for revenue productivity gains does match previous estimates.

The results for TFP gains in terms of statistical significance do not match previous estimates and the results for revenue productivity gains do match previous results. This could be because of two reasons. First, this could be due to dropping observations because of negative coefficients. Second, this multiple imputation does not work properly on the Amadeus database because some variables of some countries are missing a quite lot of information. For example, around 87 percent of data on the material costs for Hungary are missing (see Table A1 for details).

Even though the results for TFP gains do not match and the coefficients are not statistically significant, the direction of coefficients for TFP gains and revenue labor productivity gains match the previous estimates exactly. Thus, since both interaction coefficients with and without imputation are negative, it further confirms that TFP gains from international trade are higher in young economies than in old economies.

[19]The parameter estimates are summarized by taking the average over the parameter estimates from all imputed datasets. The standard errors are pooled by combining the within imputation variance and the between imputation variance. I use the following formula to calculate the standard error:

$$SE = \sqrt{Var_{within} + Var_{between} + \frac{Var_{between}}{K}}$$

where Var is variance, K is the number of imputed datasets, $Var_{within} = \frac{\sum_{i=1}^{K} SE_i^2}{K}$, $Var_{between} = \frac{\sum_{i=1}^{K}(\beta_i - \hat{\beta})^2}{K-1}$, SE is standard error, and β is the parameter estimate.

Table 10: The effect of trade openness on productivity gains

Dependent variable: TFP or revenue productivity gains from reallocating inputs								
Controls	**TFP Gains**				**Revenue Productivity Gains**			
	OLS Regressions		**IV Regressions**		**OLS Regressions**		**IV Regressions**	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Age	0.32	0.36	-0.41	-0.70	0.29	0.30	0.09	-0.21
	(0.16)	(0.16)	(0.64)	(0.71)	(0.05)	(0.05)	(0.43)	(0.43)
Openness	0.11	0.11	0.09	0.07	0.29	0.28	0.39	0.38
	(0.09)	(0.09)	(0.14)	(0.14)	(0.06)	(0.06)	(0.07)	(0.08)
Age*Openness	-0.03	-0.03	-0.03	-0.01	-0.11	-0.11	-0.15	-0.14
	(0.03)	(0.03)	(0.05)	(0.05)	(0.02)	(0.02)	(0.03)	(0.03)
Herfindahl index	-0.09	-0.11	-0.20	-0.25	-1.08	-1.09	-1.14	-1.18
	(0.44)	(0.44)	(0.56)	(0.59)	(0.12)	(0.12)	(0.14)	(0.14)
N	1876	1876	1876	1876	2068	2068	2068	2068
R squared	0.28	0.30			0.32	0.33		
Cragg-Donald F-stat			9.14	9.10			8.41	9.51

Notes: An observation is a country, a year, and an industry. The dependent variable, which is the percentage change in aggregate TFP from reallocating inputs within-industry relative to 2006, is calculated using the Dynamic Olley-Pakes Decomposition method. The age is in the log scale. Openness is measured by trade share, which is calculated by adding exports to imports and dividing the sum by outputs. See the method section for detailed explanations of the Herfindahl index and institutional quality. In all specifications, I include industry fixed effects, year fixed effects, and country fixed effects. Column (1) presents estimates without including the institutional quality and column (2) presents estimates including the institutional quality. The number of observations for TFP gains does not match previous estimates because observations are dropped due to either the labor or capital coefficient being negative. Standard errors are in parentheses. Cragg-Donald F stat is a test for weak identification.

7 Conclusions

It has been well established that international trade generates productivity gains by reallocating inputs from the least productive firms to the more productive firms. Firm age may matter for these productivity gains because it is easier and cheaper for young firms to reallocate inputs. In other words, their reallocation channel is stronger, and their productivity gains from international trade are therefore higher. Thus, I ask the question: Does increased exposure to foreign competition generate more productivity gains in relatively young economies?

The results of my study support the argument advanced by Melitz (2003) that an increase in international trade increases productivity gains by reallocating inputs within industries more efficiently. However, in addition, I find that productivity gains from reallocating inputs within industries are higher in young economies. Specifically, such gains for young economies are five times higher than the gains for old economies.

To establish this result, I estimate the effect of the interaction between mean firm age and trade share on the productivity gains from reallocating inputs within industries. To overcome the problem posed by the possible endogeneity of mean firm age, I use the IV method. As an instrument for mean firm age, I use the number of years after the birth of the current form of government. Since a change in the form of government can bring institutional changes that affect productivity gains, I add a vector for the institutional quality to the estimation.

An important question remains: What are the other probable causes for differences in productivity gains from reallocating inputs among firms across countries? Potential areas of future research would be to identify other factors that are caused for the differential effect of international trade on these productivity gains across countries.

Appendix

Table A1: Missing observations (Percentage)

Country	Value added	Labor	Capital Stock	Material Costs	Age
Bulgaria	2.55	0.23	0.19	2.55	55.53
Czech Republic	0.25	5.05	0.03	0.25	0.00
Estonia	2.64	6.84	5.79	2.64	0.00
France	2.22	55.42	0.00	2.22	0.00
Germany	54.11	15.35	0.89	54.02	0.32
Hungary	87.38	25.80	6.82	87.36	1.33
Italy	0.89	18.89	0.01	0.89	0.00
Poland	0.91	57.49	4.58	0.76	0.00
Portugal	4.94	0.47	3.51	4.90	0.05
Romania	1.57	0.02	0.00	1.52	0.00
Slovakia	0.40	6.88	0.03	0.40	0.00
Slovenia	1.54	3.31	4.49	1.38	7.97
Spain	18.96	5.08	1.86	18.88	0.02
Sweden	5.03	0.97	0.01	5.03	0.00

Note: This Table shows the percentage of information missing in value added, labor, capital, material costs, and age of firms.

Table A2: Correlation among variables

	TFP gains	Age	Openness	Age*Openness	Herfindahl index
TFP gains	1				
Age	-0.0387	1			
Openness	0.00431	-0.0956***	1		
Age*Openness	-0.00537	0.0289	0.987***	1	
Herfindahl index	-0.301***	-0.0339	0.281***	0.260***	1

Notes: TFP gains is the gains from reallocating inputs. The age is in log scale. Openness is measured by trade share. See the method section for a detailed explanation of the Herfindahl index.

Table A3: First-stage regression

	Age	Age*Openness
Instrument for Age	0.66***	0.06
	(0.10)	(0.16)
Instrument for Age*Openness	0.01	0.64***
	(0.01)	(0.02)
Openness	-0.02	0.61***
	(0.03)	(0.05)
Herfindahl index	-0.25***	-0.49***
	(0.05)	(0.07)
Institutional quality	Yes	Yes
Year	Yes	Yes
Industry	Yes	Yes
Country	Yes	Yes
N	2068	2068
R^2	0.84	0.99

Notes: An observation is a country, a year, and an industry. The age is in log scale. *Instrument* is measured as the number of years after the birth of the current form of government. Openness is measured by trade share (calculated by adding exports to imports and dividing the sum by the outputs). See the method section for detailed explanations of the Herfindahl index and institutional quality. Standard errors are in parentheses.

Table A4: The effect of trade openness on TFP gains

Controls	Robust SE				Cluster SE			
	OLS Regressions		IV Regressions		OLS Regressions		IV Regressions	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Age	0.26	0.28	-0.79	-0.91	0.26	0.28	-0.79	-0.91
	(0.07)	(0.07)	(0.38)	(0.41)	(0.14)	(0.14)	(0.46)	(0.49)
Openness	0.32	0.32	0.23	0.22	0.32	0.32	0.23	0.22
	(0.08)	(0.08)	(0.09)	(0.10)	(0.15)	(0.15)	(0.18)	(0.19)
Age*Openness	-0.10	-0.10	-0.07	-0.07	-0.10	-0.10	-0.07	-0.07
	(0.03)	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.07)	(0.07)
Herfindahl index	-0.95	-0.96	-1.16	-1.18	-0.95	-0.96	-1.16	-1.18
	(0.34)	(0.34)	(0.40)	(0.41)	(0.69)	(0.69)	(0.82)	(0.84)
Institutional quality	No	Yes	No	Yes	No	Yes	No	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1973	1973	1973	1973	1973	1973	1973	1973
R squared	.31	.32			.31	.32		
Kleibergen-Paap F-stat			18.68	24.01			19.35	28.12

Dependent variable: TFP gains from reallocating inputs

Notes: An observation is a country, a year, and an industry. The dependent variable, which is the percentage change in aggregate TFP from reallocating inputs within-industry relative to 2006, is calculated using the Dynamic Olley-Pakes Decomposition method. The age is in log scale. Openness is measured by trade share, which is calculated by adding exports to imports and dividing the sum by outputs. See the method section for detailed explanations of the Herfindahl index and institutional quality. Standard errors are in parentheses. Kleibergen-Paap F stat is a test for weak identification. I calculate the cluster standard errors at the country and industry levels.

Table A5: Percentage of industries pass the CDM test

Country	Percentage of industries pass the CDM test
Bulgaria	63.16
CzechRepublic	81.25
Estonia	55.56
France	5.88
Germany	77.78
Italy	26.32
Poland	52.63
Portugal	31.58
Romania	88.89
Slovakia	52.63
Slovenia	89.47
Spain	38.89
Sweden	52.63

Note: The table shows the percentage of industries that have passed the CDM test under significance level 0.05 by adding the size of firms, legal form, legal status, and year as covariates.

References

Ackerberg, D. A., K. Caves, and G. Frazer (2015, November). Identification Properties of Recent Production Function Estimators. *Econometrica 83*(6), 2411–2451.

Baily, M. N., C. Hulten, and D. Campbell (1992). Productivity Dynamics in Manufacturing Plants. In *Brookings Papers on Economic Activity: Microeconomics*, Volume 4, pp. 187–267. Brookings Institution.

Bartelsman, E., J. Haltiwanger, and S. Scarpetta (2009, January). Measuring and analyzing cross-country differences in firm dynamics. In T. Dunne, J. B. Jensen, and M. J. Roberts (Eds.), *Producer Dynamics: New Evidence from Micro Data*, pp. 15–76. University of Chicago Press.

Burgette, L. F. and J. P. Reiter (2010, September). Multiple Imputation for Missing Data via Sequential Regression Trees. *American Journal of Epidemiology 172*(9), 1070 – 1076.

Congleton, R. D. and D. Yoo (2014, August). A Test of the Unraveling Hypothesis: Constitutional Bargaining and the Quality of African Institutions. CEI Working Paper Series 2014-4, Center for Economic Institutions, Institute of Economic Research, Hitotsubashi University.

Feenstra, R. C., R. Inklaar, and M. P. Timmer (2015, October). The Next Generation of the Penn World Table. *American Economic Review 105*(10), 3150–82.

Fernandes, A. M. (2007, March). Trade Policy, Trade Volumes and Plant-Level Productivity in Colombian Manufacturing Industries. *Journal of International Economics 71*(1), 52–71.

Fors, H. C. and O. Olsson (2007). Endogenous Institutional Change after Independence. *European Economic Review 51*(8), 1896 – 1921.

Foster, L., J. C. Haltiwanger, and C. J. Krizan (2001, January). Aggregate Productivity Growth. Lessons from Microeconomic Evidence. In *New Developments in Productivity Analysis*, pp. 303–372. University of Chicago Press.

Griliches, Z. and H. Regev (1995, January). Firm Productivity in Israeli Industry 1979-1988. *Journal of Econometrics 65*(1), 175–203.

Haltiwanger, J., R. S. Jarmin, and J. Miranda (2013, May). Who Creates Jobs? Small versus Large versus Young. *Review of Economics and Statistics 95*(2), 347 – 361.

Jonsson, G. and A. Subramanian (2001). Dynamic Gains from Trade: Evidence from South Africa. *IMF Staff Papers 48*(1), 1–8.

Kaufmann, D. and A. Kraay (2018). Worldwide Governance Indicators.

Levinsohn, J. and A. Petrin (2003, April). Estimating Production Functions Using Inputs to Control for Unobservables. *Review of Economic Studies 70*(2), 317–341.

Li, C. (2013). Little's test of missing completely at random. *Stata Journal 13*(4), 795–809.

Lileeva, A. and D. Trefler (2010). Improved access to foreign markets raises plant-level productivity…for some plants*. *Quarterly Journal of Economics 125*(3), 1051–1099.

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association 83*(404), 1198–1202.

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

Melitz, M. J. and S. Polanec (2015, April). Dynamic Olley-Pakes Productivity Decomposition with Entry and Exit. *The RAND Journal of Economics 46*(2), 362–375.

Melitz, M. J. and D. Trefler (2012, Spring). Gains from Trade When Firms Matter. *Journal of Economic Perspectives 26*(2), 91–118.

Olley, G. S. and A. Pakes (1996, November). The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica 64*(6), 1263–97.

Pavcnik, N. (2002, January). Trade Liberalization, Exit, and Productivity Improvement: Evidence from Chilean Plants. *Review of Economic Studies 69*(1), 245–76.

Timmer, M. P., E. Dietzenbacher, B. Los, R. Stehrer, and G. J. de Vries (2015, August). An Illustrated User Guide to the World Input-Output Database: the Case of Global Automotive Production. *Review of International Economics 23*(3), 575–605.

Trefler, D. (2004, September). The long and short of the canada-u. s. free trade agreement. *American Economic Review 94*(4), 870–895.

White, T. K., J. P. Reiter, and A. Petrin (2013, January). Plant-Level Productivity and Imputation of Missing Data in the Census of Manufactures. Working Papers 11-02, Center for Economic Studies, U.S. Census Bureau.

Wooldridge, J. M. (2009, September). On Estimating Firm-Level Production Functions Using Proxy Variables to Control for Unobservables. *Economics Letters 104*(3), 112–114.