Global Retail Chains and the Supplying Industries: Evidence from Romania

Beata S. Javorcik*
Yue Li**

Published in the CESifo Economic Studies in 2014

Abstract

Despite the controversy often surrounding retail sector liberalization, there little empirical evidence on the implications of such a reform. Using data from Romania, this study sheds light on what happens to the supplying industries after a country opens its retail sector to foreign direct investment. The study relies on a unique dataset combining outlet-specific information on global retail chains with a panel of Romanian manufacturing firms. The difference-in-difference analysis finds that the expansion of global retail chains leads to a significant increase in the total factor productivity (TFP) in the supplying manufacturing industries: their entry into a region raises the TFP by 3.8 to 4.7 percent. These results suggest that liberalizing retail sector can boost performance of domestic manufacturing.

JEL classifications: F23, F14, L81, D24

Keywords: global retail chains, productivity, retail sector liberalization, foreign direct investment, backward linkages

* Department of Economics, University of Oxford and CEPR, Manor Road Building, Oxford OX1 3UQ, UK. Email: beata.javorcik@economics.ox.ac.uk.
** World Bank, 1818 H St, NW; MSN MC 10-1011; Washington, DC 20433, USA. Email: yli7@worldbank.org.
The views expressed in the paper are those of the authors and should not be attributed to the World Bank, its Executive Directors or the countries they represent.

1. Introduction

Retail sector liberalization grabbed the headlines in 2012, thanks to the controversy surrounding the issue in India. India is still a nation of small shopkeepers, where only 10% of total retail sales of $450-500 billion a year, take place in modern stores. In September of 2012, the government, led by Prime Minister Manmohan Singh, announced the reform to allow 51% foreign direct investment (FDI) in supermarkets and department stores, as a part of an unexpected and ambitious reform package. Individual states have the right to opt out from the retail sector reform, and so far only 10 states said they would allow global retail chains like Wal-Mart, Tesco and Carrefour to enter.

What can the states that have opted to open their retail sector gain and what will the states that have not forgo? As it will take time for the effects of the Indian reforms to materialize, experience of other emerging economies, which undertook similar reforms earlier, may be informative. This paper aims to provide such evidence by studying the implications of retail sector liberalization in the context of Romania. More specifically, it examines how the opening of the Romanian retail sector to FDI affected upstream manufacturing industries.

The upstream (supplying) industries can be affected by the reform in several ways. First, entry of foreign retailers may lower distribution costs, stimulate economies of scale by giving producers access to a larger market, and increase competition due to a greater ability of foreign retailers to source products from abroad. The competition effect may in turn encourage productivity improvements and innovation among suppliers. Innovation may also be encouraged by foreign retail chains serving as a conduit for information.[1]

In our analysis, we combine a unique dataset on Romanian outlets of foreign retail chains with a panel data on manufacturing firms operating in Romania during the period 1997-2005. Romania's opening to FDI is a relatively recent phenomenon, so our analysis covers both the period before and after the entry of foreign retailers. The size of the country allows us to take advantage of their uneven geographical expansion. In particular, our study takes a difference-in-difference approach, relying on the differences in the speed of the expansion of global retail chains across eight Romanian regions and the fact that only some industries within each region should be affected. Multiple indicators are used to measure the regional presence of foreign

[1] Some of these effects have been documented in a recent case study describing the effects of Wal-Mart's entry on detergent producers in Mexico (Javorcik, Keller and Tybout 2008) and in the World Bank enterprise survey analyzed in Javorcik and Li (2013).

chains, including an indicator variable, the number of outlets, and their selling space.

We find that an expansion of global retail chains leads to a significant increase in the total factor productivity (TFP) in the supplying industries in the region where the expansion took place. The presence of global retail chains in a region boosts the TFP of firms in the supplying industries by 3.8 to 4.7 percent. These conclusions hold when we use alternative measures of performance: labor productivity and the TFP measure taking into account the simultaneity bias between unobserved productivity shocks and input choices (estimated following Ackerberg, Caves and Frazer 2006). Our findings are further confirmed in specifications in first, second and long differences.

We address potential endogeneity problems in several ways. We control for unobserved heterogeneity at the region-year level. We perform a strict exogeneity test, suggested by Wooldridge (2002, p. 285), and show that it does not reject our empirical strategy. To deal with the possible omitted variable bias, we demonstrate that controlling for regional wage levels does not change the conclusions of our study.

Our paper adds to the empirical literature examining the effects of services liberalization. Arnold et al. (2010 and 2011) and Fernandes and Paunov (2012) examine the implications of services liberalization using firm-level data from India, the Czech Republic, and Chile, respectively. They find that services sector reform boosts the performance of manufacturing firms using services as inputs. The former study compares the effects of liberalizing banking, insurance, telecommunications and transport sectors, while the latter two studies conclude that allowing foreign entry into services industries is the key channel through which the effect takes place.

The literature focusing specifically on the liberalization of the retail sector is quite limited. A case study by Chavez (2002) describes the evolution of foreign retail chains and Mexican domestic retailers around the formation of NAFTA and the increasing competitive pressure caused by the entry of foreign retailers. Javorcik et al. (2008) document how the entry of Wal-Mart into Mexico facilitated the modernization of the retail sector and stimulated fundamental changes in the relationship between retailers and suppliers of soaps, detergents, and surfactants. They find that Wal-Mart's entry has driven high-cost suppliers out of business, benefited surviving producers by providing access to a larger market and prompted suppliers to introduce more innovations. In contrast, a case study by Durand (2007) concludes that FDI has

played an important role in modernizing the retail sector in Mexico, but has dampened the performance of local retailers and retail wages by introducing higher competitive pressures. The large increase in imports initiated by Wal-Mart's entry has most likely negatively affected Mexican producers in the supplying sectors. These case studies are suggestive of a strong relationship between the presence of global retail chains and the performance of supplying firms but the direction of the relationship they describe is an open question.

To the best of our knowledge, there are two causal analyses of the relationship between the entry of foreign retail chains and the performance of upstream sectors. The work by Javorcik and Li (2013), which is most closely related to this study, also focuses on Romania but addresses the issue from the perspective of economic geography. More specifically, for each of 42 Romanian counties the authors calculate a measure of access to foreign retailers. It is defined as the sum of the number of foreign outlets (or their selling space) in all counties in Romania weighted by the inverse of their distance to the county where the manufacturer in consideration operates. They then relate the TFP of manufacturing firms to the proxies for access to foreign chains and ask whether industries producing food and beverages are affected differently by the foreign chains relative to other industries. They conclude that presence of global retail chains leads to a significant increase in the productivity of these two upstream industries.

Iacovone et al. (2011) study the effects of Wal-Mart's entry into Mexico on Mexican manufacturers of consumer goods. The authors develop a dynamic industry model in which firms decide whether to sell their products through Wal-Mart or use traditional retailers. Wal-Mart provides access to a larger market but puts continuous pressure on its suppliers to improve their product's appeal and forces them to accept relatively low prices. Simulations of the model show that the arrival of Wal-Mart separates potential suppliers into two groups. Those with relatively high-appeal products choose Wal-Mart as their retailer, whereas those with lower appeal products do not. For the industry as a whole, the model predicts that the associated market share reallocations, adjustments in innovative effort, and exit patterns increase productivity and the rate of innovation. These predictions are supported by the analysis of plant-level data in regions with differing levels of Wal-Mart presence in the period 1994-2002.

Our study is structured as follows. Section 2 discusses the channels through which presence of global chains may affect supplying industries. Section 3 describes the data. The relationship between the expansion of global retailers and the productivity of the supplying

sectors is analyzed in Section 4. Conclusions are presented in the last section.

2. Data

Our paper studies the relationship between the presence of global retail chains and developments in the supplying industries in Romania. The main data source for our study is the commercial database *Amadeus* published by Bureau van Dijk, which includes data on location, contact information, industry classification, standard financial statements.

We identify global retail chains using information on company name, industry classification and ownership from *Amadeus*. We cross check *Amadeus* data against the information on major international retail chains in the *World Retail Data and Statistics 2006/2007* and *European Marketing Data and Statistics* published by the Euromonitor International, Economist Intelligence Unit (EIU) Industry Briefing *Romania: Consumer Goods and Retail Background*, the *GAIN Report* by USDA Foreign Agriculture Service and *Dun & Bradstreet Business Report*. We identify 9 foreign retail chains operating in Romania. By contacting each retail chain directly, we obtain information on the opening date of all stores, their location and selling space.[2]

Our baseline data come from the 2007 release of *Amadeus*. We employ six earlier releases (1998, 2001, 2002, 2003, 2004 and 2005) to capture firms which exited during the period studied and hence were not included in the 2007 release. Our data set is an unbalanced panel which includes new entrants and exiting firms.

We focus on establishments that report the basic information including the value of total assets, turnover and employment, or 320,373 observations on 64,767 companies. We drop observations with negative values of total assets and interpolate missing values for output and input variables. We exclude observations with unusually large fluctuations in output and input variables (i.e., we drop the top and bottom one percent of the distribution of changes in each variable). We also drop observations for years prior to 1997.[3] We also exclude observations for which the firm's age (a covariate in our model) is missing.[4] This leaves us with 250,950

[2] We did not manage to obtain this information for Kaufland, which entered Romania during the last year of the sample (and because of using lags would not have been included in our analysis anyway), and Mega Image which is one of the smaller entrants.

[3] Information for the pre-1997 period is available only for a small number of firms. From our conversations with Bureau van Dijk, we understand that this is due to the coverage of the database being extended in the late 1990s. Thus to avoid a sample selection bias, we decided to focus on the period starting in 1997.

[4] Dropping a firm's age as a covariate and keeping these observations in the sample would not affect the conclusions

observations on 52,138 firms, which is the sample used to analyze the impact on labor productivity. When we focus on the TFP as the outcome of interest, we drop observations with unusual fluctuations in the TFP measures. We end up with 221,236 observations on 49,552 companies or 245,568 observations on 51,929 companies depending on the TFP measure used.

We express output and production inputs in real terms. Output is deflated by the producer price index (PPI) specific to the 3-digit NACE sector, obtained from the *Statistical Yearbook of Romania*. Labor input is measured using the number of employees, and capital is proxied by deflated tangible fixed assets. The capital deflator is a simple average of PPI from five NACE sectors.[5] Material inputs are defined as material costs deflated by the weighted average of PPI of the supplying sectors, with the weights defined based on the input-output matrix for year 2000 provided by the Statistical Institute of Romania. Wages are deflated by the consumer price index from the IMF's *International Financial Statistics* (IFS).

In one of our robustness checks, we also include the average real wage per worker at the regional level. To calculate the average real wage, we use data on wages and employment of all companies operating in Romania during the period of 1997-2005 listed in *Amadeus* database, including all firms active in agriculture, industry, and services sectors. The data are then deflated by the consumer price index. In another specification, we also utilize information on the regional gross value added reported in the Eurostat database of regional statistics.

Finally, we also use information on imports and exports obtained from the UN's *COMTRADE* database. We concord 4-digit HS codes with the 4-digit NACE codes. We convert the data into Romanian lei using the average annual exchange rate from the IMF's IFS and deflate the figure by the GDP deflator from the World Bank's World Development Indicator.

3. Global Retail Chains in Romania and the Supplying Sectors

The expansion of global retail chains in Romania started with the entry of the German chain, Metro Cash & Carry, in 1997. It was followed by a handful of other big-box retailers, for example, Carrefour from France, REWE from Germany, and Cora from Belgium. The number of their outlets increased drastically between 1999 and 2005, from 5 to 86. Their selling spaces increased by more than ten folds to 463,000 square meters. Their employment started from

---

of this paper.

[5] These are: machinery and equipment; office, accounting, and computing machinery; electrical machinery and apparatus; motor vehicles, trailers, and semi-trailers; and other transport equipment.

virtually zero to over 18,900. By 2005, they accounted for almost a quarter of total retail sales (Figure 1).

The paths of foreign chains' expansion in Romania follow each other closely. Bucharest and its outskirts were their initial favorite location. Then, they sprawled to Western region, close to Hungary. In other words, their expansion was geographically uneven. This carries over to 2005, when their distribution was skewed toward a few regions despite the fact their outlets had presented in all eight regions.

Global retail chains are distinct in comparison with other retailers operating in Romania. They tend to have a larger workforce and a higher capital stock. The anecdotal evidence from other countries suggests that global retailers tend to be leaders in adopting advanced retail technologies, from large sales rooms and warehouses to computerized inventory tracking systems. Consistently, they are found to be more capital intensive in Romania. No surprisingly, their market shares are significantly larger. There are also more productive, as measured by sales per worker, real wage per worker, and value added per worker. But, they do not appear to be more profitable during the period under consideration, as measured by return on assets and return on sales.

What may happen to supplying industries after the entry of global retail chains? Anecdotal evidence and different strands of empirical literature suggest the performance of firms in the supplying industries may be affected in several ways. First, the entry of foreign chains may intensify competitive pressures on suppliers. Their larger size and international sourcing networks suggest that global chains may have greater bargaining power vis a vis local suppliers. This in turn forces suppliers to become more efficient.[6]

Second, global retail chains may offer access to a larger market and in this way lower the distribution costs of their suppliers. Suppliers may deliver fewer but larger shipments to several retail outlets, rather than spreading across a larger number of small mom-and-pop stores. This may allow them to produce more output with the same amount of labor and capital by saving on employee time and capital stocks (e.g. truck fleet). By allowing their suppliers to reach a larger market through their regional, national and international presence, foreign retailers may also stimulate economies of scale in the supplying industries.

[6] Additionally, if the presence of global retail chains forces local retailers to lower their prices, they are likely to pass the price pressure onto their suppliers. Thus we would expect that entry of global retail chains will affect all producers in the supplying industries and not just firms directly supplying global retailers.

Third, global retailers possess cutting-edge retail technologies and computerized inventory systems. Establishing supplying relationships with them may help suppliers to plan ahead and be better informed about changes in consumer demand.[7]

In sum, by increasing competitive pressures on suppliers, cutting distribution costs and offering easier access to information and a larger market, global retail chains may stimulate productivity growth in the supplying industries through within-firm improvements.

4. Impact on the Total Factor Productivity in the Supplying Industries

4.1 Identifying Assumptions

In our analysis of the relationship between the presence of global retail chains and the performance of the supplying industries, we take advantage of regional variation in foreign chains' expansion. We rely on the Nomenclature of Territorial Units for Statistics (NUTS) which divides Romania into eight regions with an average territory of 29,800 km^2.[8] We focus on the changes in the performance of the supplying sectors following the entry of foreign chains into their region. Our identifying assumption is that the impact of global chains' entry should be felt most strongly within the region of entry. It is justified on the grounds of producers located close to outlets of global retail chains being better informed about the type of products sold by the chains, their characteristics and pricing, and thus more likely to supply the chains.[9] Moreover, if stores are allowed to make purchasing decisions individually, they will have an incentive to minimize transport costs by sourcing from within their region.

The second identifying assumption is that industries supplying consumer products to supermarkets, as opposed to industries supplying industrial inputs, should be affected. The core products sold in all of the identified foreign supermarkets are food and beverages. Thus, we define supplying industries as food and beverage manufacturing.[10,11]

---

[7] Foster, Haltiwanger and Krizan (2006) discuss the impact of the information technology revolution on the retail trade sector. For example, adoption of systems that electronically link cash registers to scanners and credit card processing machines allows stores to track the success of their pricing strategies for individual items.

[8] The NUTS classification is a hierarchical system for dividing up the economic territory of the EU for the purpose of the collection, development and harmonization of EU regional statistics, socio-economic analyses of the regions, and framing of EU regional policies.

[9] Our identification strategy, based on geographical proximity, should work best for goods that are expensive to transport. We will examine this question in our robustness checks and show that this is indeed the case.

[10] Our list of the supplying industries includes: production, processing and preserving of meat and meat products; processing and preserving of fruit and vegetables; manufacture of dairy products; manufacture of grain mill products, starches and starch products; manufacture of other food products; manufacture of beverages. Hereafter, we refer to these industries as the food sector.

[11] We are aware of the fact that some industries producing non-food consumer products may also be affected by

All of the supplying industries are found to present in all eight NUTS regions, validating our regional approach. Manufacturing of fruit and vegetable products is represented in 37 of 42 counties in 1998 and 40 counties in 2004; manufacturing of dairy products existed in 41 counties; and the remaining four industries are spread across all 42 counties.

Plant- rather than firm-level data would be a better choice for our study. Unfortunately, plant-level information is not available to us. This means that if a firm owns plants in multiple regions, the presence of global retailers will be measured with error. However, the measurement error is likely to bias our estimates toward zero, making it more difficult to find a statistically significant effect. Thus our estimates should be thought of as a lower bound on the true effect.

4.2 Descriptive Analysis

As the first step in our analysis, we consider some descriptive statistics. We estimate the distributions of the logarithm of TFP for firms operating before and after the entry of global retail chains into their region. We do so separately for the food sector and for the remaining industries.[12] These distributions are plotted in Figure 2. We note that the distribution of productivity shifts to the right in the post-entry period in the case of the food supplying sector. The pattern for non-food industries is less clear.

The difference becomes more pronounced at the regional level. We calculate the average level of the logarithm of TFP for firms operating in a given region in a given time period. For both food and non-food sectors, we compare the distribution in the period before and after the entry of global retail chains. As shown in Figure 3, there is a clear shift of the distribution of productivity to the right in the post-entry period in the case of the food sector. The pattern for the non-food sector is not clear. While we cannot say anything about the direction of causality, these charts hint at a positive relationship between the productivity of the supplying industries and the presence of global retail chains.

As the pattern observed in Figure 2 and Figure 3 could be capturing effects of

---

entry of some (though not all) global retail chains. An earlier version of this study focused on industries manufacturing fast moving consumer goods (rather than just food) and produced qualitatively similar results. However, as we have no way of identifying non-food products sold by each chain, we choose to focus on manufactured food products. This means that our results will *underestimate* the overall effect of global retail chains on manufacturing. As in our regressions we examine how the presence of global retailers affects the performance of food industries (relative to the performance of all manufacturers), not taking into account the possibility that non-food manufacturers are also affected by global chains will *work against* us finding a statistically significant effect.

[12] As explained in detail in the next section, TFP is measured relative to a hypothetical firm operating in the same sector in the first year of the sample 1997.

macroeconomic shocks or regional trends, we proceed to examine the relationship between the expansion of global retail chains and the TFP in the food sector using a regression analysis.

4.3 Empirical Strategy

In our empirical analysis, we use a difference-in-differences approach. Namely, we compare the TFP in the supplying industries before and after the entry of foreign chains into their region with the TFP of non-supplying industries in the same region during the same period. Our analysis is based on the following specification:

$$TFP_{it} = \gamma_1 Food_s \times global\_chain_{r,t-1} + \gamma_2 age_{it} + V_{s,t-1}\Gamma + \gamma_{rt} + v_i + \mu_{it} \quad (1)$$

where TFP_{it} denotes the logarithm of the total factor productivity of manufacturing firm i's operating in sector s at time t. Our baseline measures of TFP is a multilateral index calculated following Aw, Chen and Roberts (2001). We first express an individual firm's outputs and inputs (capital, labor and materials) as deviations from a hypothetical reference firm operating in the same sector at time t with average input costs shares, average logarithm of inputs and average logarithm of outputs. Then we chain-link all reference firms together over time within a sector. The index is an extension of the multilateral TFP index derived by Caves et al. (1982), which allows for consistent comparison of TFP in firm-level data with a panel structure. Thus the index is a measure of the proportional difference in TFP for firm i observed in year t relative to the hypothetical firm in the base year, 1997. The index is calculated separately for each of the 3-digit NACE manufacturing sectors. The advantage of this measure is that it allows for heterogeneity in production technology across producers. The disadvantage is that it does not take into account economies of scale and does not allow for measurement error.

As a robustness check, we also employ two alternative measures of firm performance. The first measure is the TFP estimated following Ackerberg, Caves and Frazer (2006) who build on the earlier work of Olley and Pakes (1996) and Levinsohn and Petrin (2003). This measure allows us to take into account the possibility that a firm's private knowledge of its productivity (unobserved by the econometrician) may affect the input decisions. It also provides a solution to multicollinearity issues from which the estimators proposed in these earlier papers may suffer. In contrast to the TFP index, the measure allows for measurement error, but does not allow production technology to differ between firms within the same industry.[13] The second measure

[13] We are grateful to Carolina Villegas-Sanchez for sharing with us a STATA routine implementing the procedure

employed is the labor productivity, defined as the value added per worker. All three measures lead to qualitatively the same conclusions. We choose the TFP index as our baseline measure as it allows for flexible and heterogeneous production technology within sectors.[14]

The explanatory variable of interest is the interaction term between the dummy for food supplying industries, denoted as $Food_s$, and a measure capturing the presence of global retail chains in the NUTS region, denoted as $global\_chain_{r,t-1}$. As explained above, we narrowly define the supplying industries as sectors manufacturing food and beverage products.

To take advantage of regional variation in their entry, we quantify the presence of global chains in seven different ways. Our first measure is a dummy taking on the value of 1 if at least one global retail chain is present in the region r at time t, and zero otherwise.[15] As our second measure, we use the number of global retail chain outlets in the region r at time t in a logarithmic form, adding one before taking a log. The third measure is the logarithm of the chains' total selling space in the region at time t (again one is added before taking a log).[16] The next two proxies are designed to measure the presence of foreign chains relative to the size of the food manufacturing sector. They are defined as the number of outlets or the total selling space divided by the total output of food manufacturers operating in a given region in a given time period. The final two measures express the presence of foreign chains relative to the economic size of the region. Thus the number of outlets or the total selling space are normalized by the gross value added of the region at time t. We lag all the measures by one period to take into account the time lag needed for the effect to manifest itself and to attenuate potential endogeneity problems.

We also control for other factors that may affect the performance of manufacturing firms. We use the number of years since establishment of a manufacturer to control for learning-by-doing effects. The variable is denoted as age_{it} and enters in a logarithmic form. We control for the effects of trade liberalization by including sector-level imports and exports.[17] Both variables are lagged one period and take a logarithmic form. The level of competition in the industry is another potential factor influencing firm productivity, and we use the Herfindahl index to take it

---

proposed by Ackerberg et al. (2006).

[14] A comparison of the robustness of five widely used productivity measures (index numbers, data envelopment analysis, stochastic frontiers, GMM and semiparametric estimation) suggests that the approach we chose tends to produce very robust results across the different experiments (see van Biesebroeck 2007).

[15] If the entry took place in the last quarter of the year, we consider it effective as of the following year.

[16] In this case one corresponds to a one square meter of selling space.

[17] Although normalizing imports or exports by domestic production may be prefereable, the differences between the classifications in which the two types of data are available means that the ratios could be misleading.

into account. Summary statistics for all variables are listed in Table 1.

To take into account the uneven economic development across Romanian regions, we control for time-varying regional factors by including a set of region-year fixed effects. We also include firm fixed effects to take into account unobservable firm characteristics, such as managerial ability. These fixed effects will also allow us control for time-invariant sector characteristics, for instance, the level of development in the pre-transition period and extent of privatization during the early reform period.

Finally, we correct the standard errors to take into account the fact that the measures of global retail chains' presence are at the region-year level while the dependent variable is at the firm-year level. Failure to correct for such data structure may lead to a downward bias in the estimated errors. We perform the correction by clustering standard errors at the region-year level.

4.4 Baseline Results

In our baseline results, reported in Table 2, the firm performance is measured using the multilateral TFP index. We present the estimates for each of the seven measures of global chain presence separately and report results of specifications without time-variant sector-specific controls together with the full model. Note that we do not need to include the variable $global\_chain_{r,t-1}$ by itself in the model because productivity changes coinciding with the chain's entry and affecting all manufacturing sectors equally are captured by region-year fixed effects.

We find that expansion of global retail chain leads to a significant increase in the total factor productivity of the food sector. This effect is statistically significant at the one percent level across all specifications. The presence of foreign chains increases the TFP of firms in the food sector located in the region by 3.8 to 4.7 percent (based on the results from column 1 and 2). The results from column 3 and 4 indicate that doubling the number of chains will lead to a 3.3 to 3.7 percent increase in firm productivity among food suppliers. If we take as a benchmark the average regional growth rate of the number of foreign chains' outlets (50 percent), the model predicts that the TFP of food suppliers increases by 1.65 to 1.85 percent per year in a region where foreign chains expand at the average speed. The results in column 5 and 6 suggest that doubling the selling space increases the TFP of food supplying sectors by about 0.4 to 0.5 percent. As illustrated in the remaining four columns, the results are robust to normalizing the measures of chain presence by either the value of output in the region's food sector or the

region's gross value added.

The productivity effects in Table 2 are comparable in magnitude to those found by other studies examining the effects of openness to FDI on firm performance. For instance, using Lithuanian data, Javorcik (2004) finds that doubling the foreign presence in downstream manufacturing sectors is associated with a 3.8 percent rise in the TFP of domestic firms in the supplying industries. Arnold et al. (2011) show that a one-standard-deviation increase in foreign presence in services industries is associated with a 7.7 percent increase in the productivity of manufacturing firms relying on services inputs.

As for the control variables, the coefficient on firm age is positive and significant across all specifications, which is consistent with learning-by-doing effects. The Herfindahl index bears a negative and statistically significant coefficient. It suggests that higher concentration is correlated with a lower productivity, which is in line with the belief that more competition encourages better performance. Imports are negatively correlated with firm productivity and exports do not appear to matter. The results on imports differ from the conclusions of Pavcnik (2002) for Chile and Fernandes (2006) for Colombia, but are in line with the findings of Arnold et al. (2011) for the Czech Republic.

4.5 Alternative Measures of Performance

We employ an alternative measure of the TFP, suggested by Ackerberg et al. (2006), which corrects for the simultaneity bias between productivity shocks and input choices. This alternative TFP measure is highly correlated with the TFP index used in the baseline results (correlation of 0.72). As before, we find that the entry of global retail chain leads to a significant increase in the TFP of the food supplying sectors (see the top panel of Table 3). The estimated coefficients are statistically significant at the one percent level in all specifications. Doubling the number of foreign chains' outlets in the region increases the TFP of the food sector in the region by 5.5 to 6 percent.

Next, we focus on labor productivity as the outcome variable (see the lower panel of Table 3). As labor productivity is strongly affected by the capital intensity of the production process, we control for the log of the firm's capital-labor ratio. Using this alternative measure of firm performance reinforces our earlier findings. We find a positive link between the expansion of foreign retail chains and the performance of food producers in the region. The estimated

coefficients are significant at the one percent level in all specifications. As expected, we also find that capital intensity is positively correlated with labor productivity.

As all three measures of firm performance lead to the same conclusions, in what follows we only report the results for the TFP index. However, using the alternative measures would not change the conclusions of the study. To save space, we will also restrict our attention to the first three measures of foreign chain presence.

4.6 Robustness Checks

We subject our results to a series robustness checks. First, we consider a possible outlier problem. Bucharest as the capital of Romania has disproportionate concentration of economic activity and wealth. It produces about twenty percent of the country's GDP while only accounting for ten percent of the total population. To check whether our results are affected by the special case of Bucharest, we exclude firms operating in Bucharest and perform the baseline analysis. All coefficients on the presence of global chains remain positive and significant at the one percent level. The magnitudes are very similar to those found in Table 2. Thus these estimates (reported in the top panel of Table 4) indicate that our findings are not driven by the observations from Bucharest.

Second, if our identification strategy based on the regional variation is reasonable, we would expect to observe that goods that are more expensive to transport should be more affected by the entry of global retail chains into the region because they are more dependent on the regional retail infrastructure than other products. To check whether this is true we use the data on the product-specific transport costs from Javorcik and Narciso (2008).[18] Each of 28 4-digit NACE sectors producing food is classified as facing high transport costs if its costs are above the median value found for all food industries. We augment our baseline specification by adding an interaction terms between the dummy for high transport cost food industries and the presence of local chains. If food industries facing high transport costs are affected more by the entry of global retailers than the food sector in general, then this additional interaction term will be

[18] The figures are derived from the difference between the value of Finnish imports from Germany recorded including the cost of insurance and freight (c.i.f. basis) and the value of German exports to Finland recorded in Germany net of the cost of insurance and freight (f.o.b. basis). Finland was chosen by Javorcik and Narciso (2008) as the importing country because it was ranked by Transparency International as the least corrupt country in the world, which suggests that Finnish import data are unlikely to be contaminated by tariff evasion. Annual figures are available in 6-digit HS classification for 1992-2005. We compute the average value for each HS code and then concord the HS classification with 4-digit NACE codes.

positive and statistically significant. This is indeed the case in all specifications: the interaction term is statistically significant at the one percent level. The interaction term between *Food* and the presence of foreign chains in the region also remains positive and statistically significant at the one percent level in all models. We conclude that these results (reported in the middle panel of Table 4) give credibility to our identification strategy and confirm our findings on the link between the activities of foreign retail chains and the firm performance in the supplying industries.

Third, to make sure that our results are not driven by mismeasurement of capital stock driven by differences in valuation of capital between old and new firms, we check whether young food producers (defined as firms in the first three years of their operation) are affected differently. Another justification for this exercise comes from the observation made by Foster, Haltiwanger and Syverson (2008) that young producer tend to charge lower prices than incumbents and thus their productivity advantage may be understated. This does not appear to be the case. The additional interaction term is statistically significant only in two of six specifications, and the variable of interest is not affected by this modification. See the bottom panel of Table 4.

Fourth, Katayama et al. (2009) argue that the problems with using revenue based TFP measures are reduced in a differenced specification, we check whether our results are robust to differencing. In the specification in first and second differences, we drop firm age but still include region-year fixed effects. We also cluster standard errors at the region-year level. We present the results in Table 5. The interaction term between *Food* and each of the three measures of the regional presence of global chains remains positive and statistically significant in all specifications. The magnitude of the impact is smaller than in the baseline specification when the presence of foreign chains is measured as a dummy or in terms of their selling space. When we use the number of outlets to proxy for foreign chains' regional presence, the magnitudes are similar to the baseline results. In summary, our conclusions remain robust.

We also conduct a simple cross-sectional regression on the overall changes in TFP between 1997 and 2005. The measures of regional presence of global retail chains and sector-specific variables are lagged by one period and span 1996 to 2004. Correspondingly, we only include region fixed effects and cluster the standard errors at the region level. The results are presented in the bottom panel of Table 5. The overall changes in the TFP of the food supplying

sectors during the period are shown to be positively correlated with changes in the regional presence of global chains. The estimated coefficients are statistically significant at the one percent level in all specification.

Finally, we examine whether our results are not subject to autocorrelation problem when using dummy for the presence of foreign chains. Bertrand et al. (2004) show that estimations with a difference-in-difference method using panel data are likely to be subject to serial correlation problems and the standard errors could be severely underestimated. To check for this potential estimation bias, we take their advice and ignore the time-series information when computing standard errors.[19] As evident from Table 6, in both specifications the variable of interest remains positive and significant at the one percent level though its magnitude becomes smaller. Given these findings, we feel reasonably confident that our baseline results are not subject to the autocorrelation problem.

4.7 Potential Endogeneity Problem and Omitted Variables Bias

To address a potential endogeneity problem, we check whether there is evidence of an impact *before* the actual entry of global chains takes place in the region. As economic conditions vary across regions, global retail chains may choose to operate in regions where food suppliers are highly productive in the first place. If such reverse causality exists, food suppliers in regions that attract global chains should exhibit higher TFP before the entry of global chains. We are not very concerned about this possibility because our specification controls for region-year fixed effects. Nevertheless, to capture firm performance in the pre-entry period, we define a new variable which takes the value of one in the year *prior* to the entry of global chains into the region, and zero otherwise. We include an interaction term between *Food* and this new dummy in our estimation. We report the results in Table 7. The new interaction term does not appear to matter, while the interactive term between *Food* and global chain presence remains positive and statistically significant. We find that the magnitudes of the coefficients on these two variables are significantly different from each other. These findings suggest that global retail chains are not

[19] We perform the test in three steps. First, we regress the logarithm of TFP on control variables (other than the variable of interest) and fixed effects and keep the residuals for food supplying sectors. Second, we divide the residuals into two groups: residuals from the years before foreign chains' entry and residuals from post-entry period and calculate a within-firm average for each period. Finally, we regress the two-period panel of mean residuals on the dummy denoting the presence of global retail chains. We also repeat the procedure for a one-period lag of the of the foreign chains' presence.

attracted to regions with more productive food producers and thus give us confidence that reverse causality is unlikely to be a serious problem in our analysis.

To further investigate the possibility that our results are driven by reverse causality we conduct a strict exogeneity test. As suggested by Wooldridge (2002, p.285), we estimate a specification with a contemporaneous interaction between *Food* and global chain presence as well as its lead and lag. We find a positive and significant effect of the lagged term in all six specifications (see Table 8). The contemporaneous effect is positive in all models, but it does not reach the conventional significance levels. More importantly, we find that the lead term bears a negative coefficient and is not significant in any of the models. These findings indicate that our analysis does not suffer from the reverse causality problem.

As for omitted variables, one may be concerned that an increase in the regional income stimulates regional demand for food industries more than that for other industries and that a higher demand encourages food production and at the same time attracts global retail chains to the region. To examine our results against this alternative explanation, we compute the average real wage in the region and use it as a proxy for the regional income level. We add an interaction term between *Food* and the log of the average wage in the region to our model. The results, reported in Table 9, show that our findings are robust to this modification. The additional interaction term does not appear to be statistically significant in any of the models and bears a negative sign. This pattern suggests that the relationship between regional income level and productivity does not exhibit systematic differences across food and non-food sectors (note that the effect of regional wages on all industries is captured by region-year fixed effects). Our variable of interest, the interaction term between *Food* and the global chain presence, remains positive and statistically significant in all of the specifications. It suggests that the regional expansion of global retail chains facilitates productivity growth of food industries located in the same region.

In sum, the above robustness checks make us confident that our conclusions are not driven by reverse causality or omitted variables.

5. Conclusions

Retail sector liberalization is a controversial issue. As shown by the recent experience of India, it is often met with hostility from domestic interests. This hostility may in part be due to the limited empirical evidence on the implications of the retail sector reform.

This paper contributes to the policy debate by presenting evidence from Romania which only recently witnessed entry of global retail chains. It focuses on the implications for the upstream manufacturing industries which distribute their products through the retail sector. Exploring variation in the timing of entry of foreign retailers into Romanian regions and the fact that only some industries would be affected by their presence, this study performs a difference-in-differences analysis. The econometric results lead us to the following conclusions. The expansion of global retail chains leads to a significant increase in the total factor productivity in the upstream industries. The chains' presence in a region increases TFP of firms in the upstream industries by 3.8 to 4.7 percent and doubling the number of chains leads to a 3.3 to 3.7 percent increase. These results suggest opening the retail sector to foreign investors can boost the performance of domestic manufacturing.

References

Ackerberg, Daniel, Kevin Caves, and Garth Frazer (2006). Structural Identification of Production Functions. Mimeo, UCLA

Arnold, Jens, Beata S. Javorcik and Aaditya Mattoo (2011). Does Service Benefit Manufacturing Firms? Evidence from the Czech Republic. *Journal of International Economics* 85(1): 136-146

Arnold, Javorcik, Lipscomb and Mattoo (2010). Services Reform and Manufacturing Performance: Evidence from India. CEPR Discussion Paper 8011

Aw, Yan Bee, Xiaomin Chen and Mark J. Roberts (2001). Firm-level Evidence on Productivity Differentials and Turnover in Taiwanese Manufacturing. *Journal of Development Economics* 66: 51-86

Bertrand, Marianne, Esther Duflo and Sendhil Mullainathan (2004). How Much Should We Trust Difference-in-Differences Estimates? *The Quarterly Journal of Economics* 2004: 249-275

Caves, Douglas W., Laurits R. Christensen and W. Erwin Diewert (1982). Multilateral Comparisons of Output, Input and Productivity Using Superlative Index Numbers. *The Economic Journal* 92: 73-86

Chavez, Manuel (2002). The Transformation of Mexican Retailing with Nafta. *Development Policy Review*, 20(4): 503-13

Dun & Bradstreet. *Dun & Bradstreet Business Report*

Durand, Cedric (2007). Externalities from Foreign Direct Investment in the Mexican Retailing Sector. *Cambridge Journal of Economics* 31 (3): 393-411

Economist Intelligence Unit. Country Profile: Romania. Various years

Economist Intelligence Unit. Industry Briefing, Romania: Consumer Goods and Retail Background. Sep. 2002, Dec. 2003, Dec. 2004, Dec. 2005, Dec. 2006

Fernandes, Ana and Caroline Paunov (2012). Foreign direct investment in services and manufacturing productivity: Evidence for Chile. *Journal of Development Economics* 97(2): 305-321

Foster, Lucia, John Haltiwanger and C.J. Krizan (2006). Market Selection, Reallocation, and Restructuring in the U.S. Retail Trade Sector in the 1990s. *The Review of Economics and Statistics* 88(4): 748–758

Foster, Lucia, John Haltiwanger and Chad Syverson (2008). Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability? *American Economic Review* 98(1): 394-425

Iacovone, Leonardo, Beata S. Javorcik, Wolfgang Keller and James Tybout (2011). Wal-Mart in Mexico: The Impact of FDI on Innovation and Industry Productivity. NBER Working Paper No. 17204

Javorcik, Beata S. (2004). Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers through Backward Linkages. *American Economic Review* 94(3): 605-627

Javorcik, Beata S. and Yue Li (2013) Do the Biggest Aisles Serve a Brighter Future? Global Retail Chains and Their Implications for Romania. *Journal of International* Economics, 90(2): 348–363

Javorcik, Beata S. and Gaia Narciso (2008). Differentiated Products and Evasion of Import Tariffs. *Journal of International Economics* 76(2): 208-222

Javorcik, Beata S., Wolfgang Keller and James R.Tybout (2008). Openness and Industrial Response in a Wal-Mart World: A Case Study of Mexican Soaps, Detergents and Surfactant Producers. *The World Economy*, 31(12)

Katayama, Haijime, Shihua Lu and James R.Tybout (2009). Firm-level Productivity Studies: Illusions and a Solution. *International Journal of Industrial Organization* 27(3): 403-413

Levinsohn, James and Amil Petrin (2003). Estimating Production Functions Using Inputs to Control for Unobservables. *Review of Economic Studies* 70(2): 317-341

Olley, Steven and Ariel Pakes (1996). The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica* 64:1263-1295

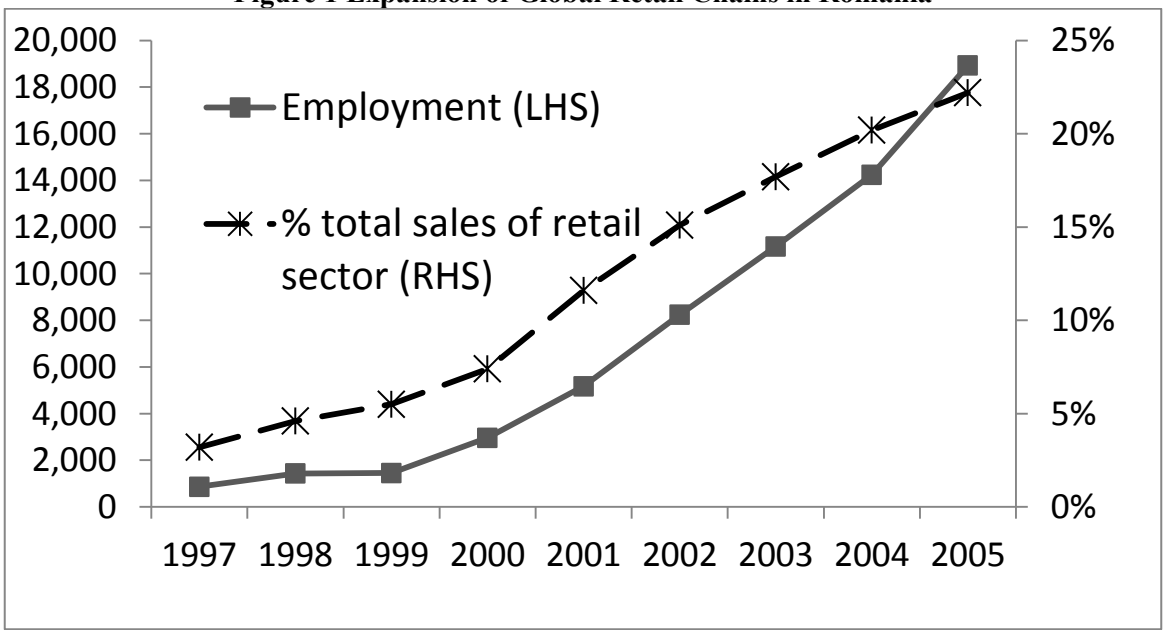
Pavcnik, Nina (2002). Trade liberalization, Exit, and Productivity improvements: Evidence from Chilean Plants. *Review of Economic Studies* 69: 245-76

Statistical Institute of Romania. *Statistical Yearbook of Romania*

Van Biesebroeck, Johannes (2007). Robustness of Productivity Estimates. *Journal of Industrial Economics* LV(3): 529-569

Wooldridge, Jeffrey (2002). *Economic Analysis of Cross Section and Panel Data*, Cambridge MA and London: The MIT Press.

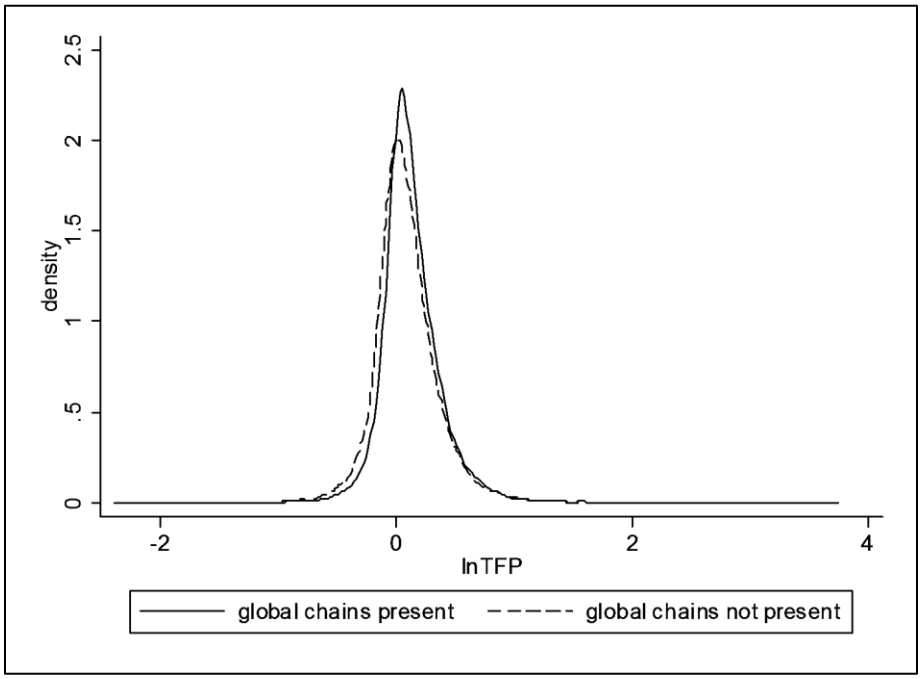
Figure 1 Expansion of Global Retail Chains in Romania



Source: Authors' calculations based on the *Amadeus* database.

Figure 2 Logarithm of Total Factor Productivity of Manufacturing Firm Pre- vs. Post-entry of Global Chains, Firm-Level Data

Food Sector



Other Industries

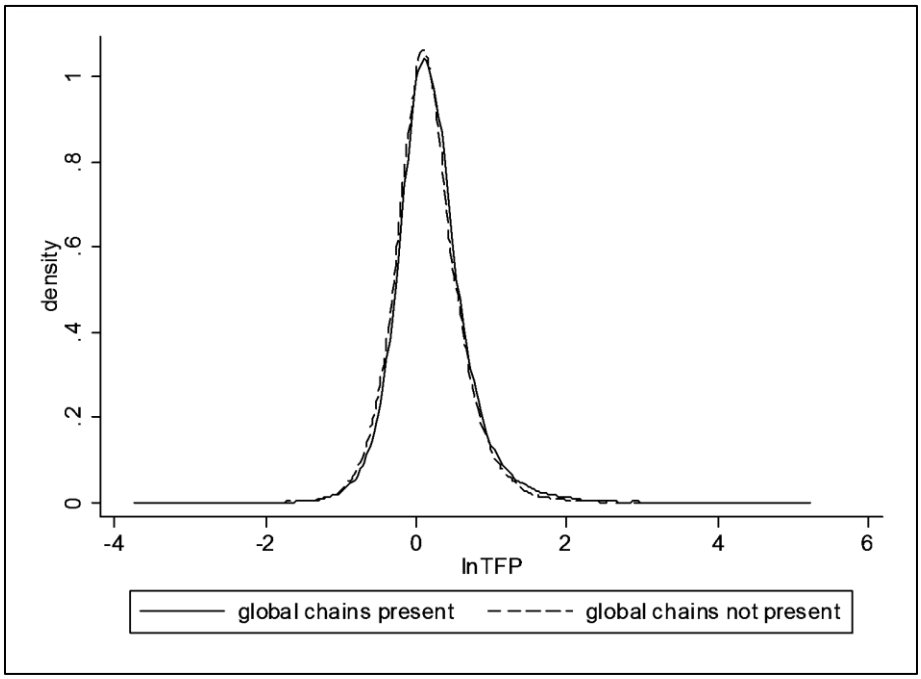
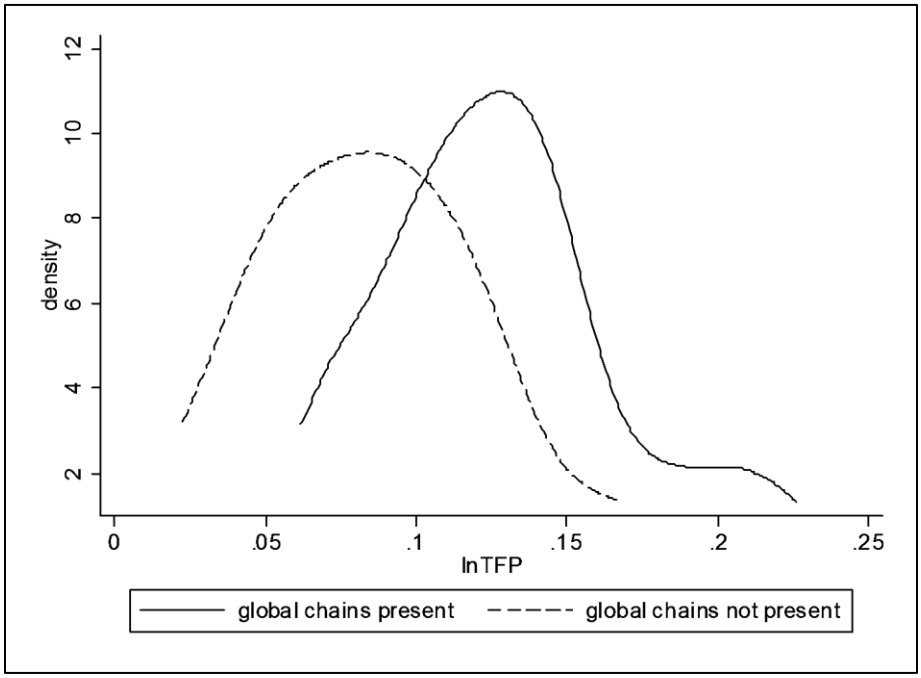


Figure 3 Logarithm of Total Factor Productivity of Manufacturing Firms

Pre- vs. Post-entry of Global Chains, Regional Average

Food Sector



Other Industries

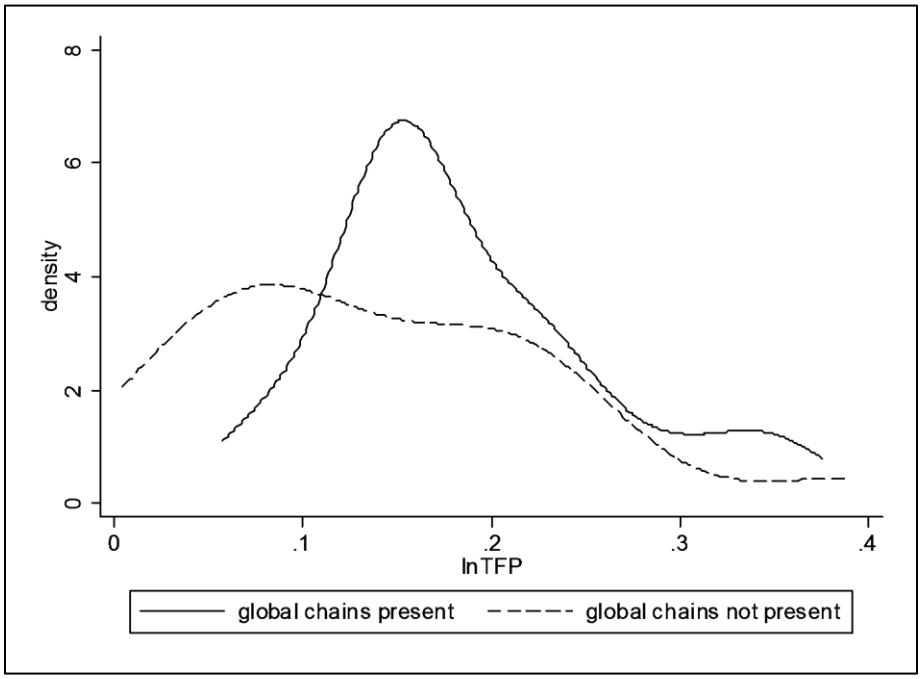


Table 1 Summary Statistics*

Firm-specific

Variable	Food Sector			Non-Food Sector		
	No. of Obs.	Mean	Std. Dev.	No. of Obs.	Mean	Std. Dev.
TFP index	49498	0.116	0.279	171297	0.188	0.491
TFP (Ackerberg et al.)	56772	0.356	0.857	188796	1.025	0.977
value added/labor (th lei 2000)	58252	7.9	29.2	192698	13.2	39.9
capital stock/labor (th lei 2000)	58252	10.8	76.7	192698	10.4	74.1
output (th lei 2000)	58252	1041.3	7388.7	192698	1529.8	20328.3
wage costs (th lei 2000)	57234	85.2	604.5	188869	267.8	2442.5
material costs (th lei 2000)	58252	739.6	5090.6	192698	897.9	13812.3
capital stock (th lei 2000)	58252	354.5	3531.0	192698	694.7	13469.9
employment	58252	23.3	155.4	192698	56.8	343.1
age	58252	7.7	3.4	192698	7.3	3.6

NACE sector-specific

Variable	Food Sector			Non-Food Sector		
	No. of Obs.	Mean	Std. Dev.	No. of Obs.	Mean	Std. Dev.
Herfindahl Index	54	0.031	0.017	808	0.165	0.179
imports (th lei 2000)	54	169032.3	174615.6	793	628703.0	1006053.0
exports (th lei 2000)	54	48700.4	55999.3	793	565333.2	1466706.0

NUTS region-specific

Variable	No. of Obs.	Mean	Std. Dev.
wage/L (th lei 2000)	72	4.6	1.0
foreign chains present (dummy)	72	0.7	0.5
no. of foreign chain outlets	72	3.8	4.3
selling space of foreign chain outlets (sq. meters)	72	22273	24194
no. of foreign chain outlets / regional food output (per mn lei 2000)	72	0.0052	0.0056
selling space of foreign chains/ regional food output (sq. meters per mn lei 2000)	72	21.07	19.07
no. of foreign chain outlets / regional gross value added (per mn lei 2000)	64	0.0004	0.0003
selling space of foreign chains / regional gross value added (sq. m per mn lei 2000)	64	1.74	1.74

*The average exchange rate in year 2000 was 2.17 Romanian Lei per US dollar. Thus the average firm in the food sector had output worth 480 thousand US dollars and capital stock of 163 thousand US dollars. The corresponding figures for the non-food sector were 705 and 320 thousand dollars, respectively.

Table 2 Presence of Global Retailers and Performance of Food Producers
Fixed effects, TFP index

	chains present		ln(no. of outlets)		ln(selling space)		Normalized by the output of food sector in the region				Normalized by gross value added of the region			
							ln(no. of outlets)		ln(selling space)		ln(no. of outlets)		ln(selling space)	
$Food_s$	0.047***	0.038***	0.037***	0.033***	0.005***	0.004***	0.044***	0.040***	0.005***	0.004***	0.039***	0.035***	0.005***	0.004***
$*(global\_chain)_{r,t-1}$	(0.010)	(0.010)	(0.005)	(0.005)	(0.001)	(0.001)	(0.007)	(0.006)	(0.001)	(0.001)	(0.006)	(0.006)	(0.001)	(0.001)
age_{it}	0.118***	0.117***	0.119***	0.118***	0.118***	0.117***	0.119***	0.118***	0.118***	0.117***	0.119***	0.118***	0.118***	0.117***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
$ln(imports)_{s,t-1}$		-0.029***		-0.020***		-0.027***		-0.021***		-0.027***		-0.020***		-0.027***
		(0.005)		(0.005)		(0.005)		(0.005)		(0.005)		(0.005)		(0.005)
$ln(exports)_{s,t-1}$		-0.004		-0.004		-0.004		-0.004		-0.004		-0.004		-0.004
		(0.004)		(0.004)		(0.004)		(0.004)		(0.004)		(0.004)		(0.004)
$Herfindahl\ Index_{st}$		-0.190***		-0.210***		-0.196***		-0.208***		-0.195***		-0.211***		-0.195***
		(0.040)		(0.039)		(0.040)		(0.040)		(0.040)		(0.039)		(0.040)
R-squared	0.019	0.02	0.02	0.021	0.019	0.02	0.02	0.021	0.019	0.02	0.02	0.021	0.019	0.02
No. of obs.	221236	220002	221236	220002	221236	220002	221236	220002	221236	220002	221236	220002	221236	220002
No. of firms	49552	49390	49552	49390	49552	49390	49552	49390	49552	49390	49552	49390	49552	49390

All models include firm fixed effects and region-year fixed effects. Standard errors, clustered at the region-year level, are reported in parentheses.
* significant at 10%, ** at 5%, *** at 1%

Table 3 Presence of Global Retailers and Performance of Food Producers
Fixed effects, TFP estimated following Ackerberg, Caves and Frazer (2006) and Labor Productivity

	chains present		ln(no. of outlets)		ln(selling space)		Normalized by the output of food sector in the region				Normalized by gross value added of the region			
							ln(no. of outlets)		ln(selling space)		ln(no. of outlets)		ln(selling space)	
ln(TFP ACF)														
$Food_s$	0.089***	0.076***	0.060***	0.055***	0.010***	0.008***	0.076***	0.071***	0.010***	0.009***	0.065***	0.060***	0.010***	0.008***
$*(global\_chain)_{r,t-1}$	(0.028)	(0.026)	(0.014)	(0.013)	(0.003)	(0.003)	(0.017)	(0.016)	(0.003)	(0.003)	(0.016)	(0.015)	(0.003)	(0.003)
Industry controls	no	yes	no	yes	no	yes	no	yes	no	yes	no	yes	no	yes
R-squared	0.02	0.022	0.021	0.022	0.021	0.022	0.021	0.022	0.021	0.022	0.021	0.022	0.021	0.022
No. of obs.	245568	244197	245568	244197	245568	244197	245568	244197	245568	244197	245568	244197	245568	244197
No. of firms	51929	51764	51929	51764	51929	51764	51929	51764	51929	51764	51929	51764	51929	51764
ln(Labor productivity)														
$Food_s$	0.081***	0.066***	0.055***	0.049***	0.009***	0.007***	0.070***	0.062***	0.009***	0.008***	0.061***	0.053***	0.009***	0.007***
$*(global\_chain)_{r,t-1}$	(0.025)	(0.024)	(0.013)	(0.012)	(0.002)	(0.002)	(0.016)	(0.015)	(0.003)	(0.002)	(0.015)	(0.014)	(0.003)	(0.002)
$ln(capital/labor)_{it}$	0.327***	0.328***	0.327***	0.328***	0.327***	0.328***	0.327***	0.328***	0.327***	0.328***	0.327***	0.328***	0.327***	0.328***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Industry controls	no	yes	no	yes	no	yes	no	yes	no	yes	no	yes	no	yes
R-squared	0.108	0.109	0.108	0.109	0.108	0.109	0.108	0.109	0.108	0.109	0.108	0.109	0.108	0.109
No. of obs.	250950	249557	250950	249557	250950	249557	250950	249557	250950	249557	250950	249557	250950	249557
No. of firms	52081	51916	52081	51916	52081	51916	52081	51916	52081	51916	52081	51916	52081	51916

All models include a firm's age, firm fixed effects and region-year fixed effects. Industry-level controls include: $\ln(imports)_{t-1}$, $\ln(exports)_{t-1}$ and Herfindahl $index_t$.
Standard errors, clustered at the region-year level, are reported in parentheses.
* significant at 10%, ** at 5%, *** at 1%

Table 4 Presence of Global Retailers and Performance of Food Producers, Robustness Checks
Fixed effects, TFP index

	chains present		ln(no. of outlets)		ln(selling space)	
Excluding Bucharest						
$Food_s*(global\_chain)_{r,t-1}$	0.044***	0.036***	0.037***	0.034***	0.005***	0.004***
	(0.010)	(0.009)	(0.006)	(0.005)	(0.001)	(0.001)
Industry controls	no	yes	no	yes	no	yes
R-squared	0.02	0.021	0.021	0.022	0.02	0.021
No. of obs.	186892	185845	186892	185845	186892	185845
High vs. low transport costs						
$Food_s*(global\_chain)_{r,t-1}$	0.035***	0.028***	0.030***	0.028***	0.004***	0.003***
	(0.010)	(0.009)	(0.005)	(0.005)	(0.001)	(0.001)
$High\ cost_s*Food_s*(global\_chain)_{r,t-1}$	0.028***	0.025***	0.016***	0.013***	0.003***	0.003***
	(0.007)	(0.008)	(0.004)	(0.004)	(0.001)	(0.001)
Industry controls	no	yes	no	yes	no	yes
R-squared	0.019	0.02	0.02	0.021	0.02	0.02
No. of obs.	221236	220002	221236	220002	221236	220002
Young vs. old firms						
$Food_s*(global\_chain)_{r,t-1}$	0.047***	0.038***	0.037***	0.033***	0.005***	0.004***
	(0.010)	(0.010)	(0.005)	(0.005)	(0.001)	(0.001)
$Young*Food_s*(global\_chain)_{r,t-1}$	0.001	0.007	0.015***	0.016***	0.000	0.001
	(0.007)	(0.007)	(0.004)	(0.004)	(0.001)	(0.001)
Industry controls	no	yes	no	yes	no	yes
R-squared	0.019	0.02	0.02	0.021	0.019	0.02
No. of obs.	221236	220002	221236	220002	221236	220002

All models include a firm's age, firm fixed effects and region-year fixed effects. Industry-level controls include: $\ln(imports)_{t-1}$, $\ln(exports)_{t-1}$ and Herfindahl index$_t$. Standard errors, clustered at the region-year level, are reported in parentheses.* significant at 10%, ** at 5%, *** at 1%

Table 5 Presence of Global Retailers and Performance of Food Producers, Specifications in Differences, TFP index

	chains present		ln(no. of outlets)		ln(selling space)	
First differences						
$Food_s * \Delta(global\_chain)_{r,t-1}$	0.021**	0.022***	0.034***	0.036***	0.003***	0.003***
	(0.008)	(0.008)	(0.008)	(0.007)	(0.001)	(0.001)
Industry controls	no	yes	no	yes	no	yes
R-squared	0.014	0.014	0.014	0.015	0.014	0.014
No. of obs.	164668	163623	164668	163623	164668	163623
Second differences						
$Food_s * \Delta(global\_chain)_{r,t-1}$	0.026*	0.028**	0.031***	0.035***	0.003**	0.004**
	(0.014)	(0.014)	(0.007)	(0.008)	(0.001)	(0.001)
Industry controls	no	yes	no	yes	no	yes
R-squared	0.014	0.014	0.014	0.015	0.014	0.014
No. of obs.	130181	129356	130181	129356	130181	129356
Long differences						
$Food_s * \Delta(global\_chain)_{r,t-1}$	0.172***	0.164***	0.074***	0.066***	0.016***	0.015***
	(0.012)	(0.015)	(0.008)	(0.007)	(0.001)	(0.001)
Industry controls	no	yes	no	yes	no	yes
R-squared	0.021	0.022	0.019	0.021	0.021	0.022
No. of obs.	11418	11346	11418	11346	11418	11346

Industry-level controls include: $\Delta\ln(\text{imports})_{t-1}$, $\Delta\ln(\text{exports})_{t-1}$ and ΔHerfindahl index $_t$.
All models in first and second differences include region-year fixed effects. Standard errors, clustered at the region-year level, are reported in parentheses.
Models in long difference include region fixed effects. The standard errors are clustered at the region level.
* significant at 10%, ** at 5%, *** at 1%

Table 6 Robustness Check on Autocorrelation
TFP index

First Stage Estimation

ln(firm age) $_{it}$	0.117***	0.117***
	(0.004)	(0.004)
ln(imports) $_{s,t-1}$	-0.038***	-0.038***
	(0.003)	(0.003)
ln(exports) $_{s,t-1}$	-0.004**	-0.004**
	(0.002)	(0.002)
Herfindahl Index $_{st}$	-0.152***	-0.152***
	(0.032)	(0.032)
R-squared	0.02	0.02
No. of obs.	220002	220002

Second Stage Estimation, only food produers

(global_chain) $_{r,t}$	0.028***	
	(0.002)	
(global_chain) $_{r, t-1}$		0.024***
		(0.002)
R-squared	0.012	0.009
No. of obs.	14775	14588

First stage estimation includes firm fixed effects and region-year fixed effects.
Standard errors are reported in parentheses.
* significant at 10%, ** at 5%, *** at 1%

Table 7 Pre-entry Impact
Fixed effects, TFP index

	chains present		**ln(no. of outlets)**		**ln(selling space)**	
Food $_s$(global_chain) $_{r,t-1}$*	0.045***	0.035***	0.036***	0.032***	0.005***	0.004***
	(0.012)	(0.011)	(0.006)	(0.006)	(0.001)	(0.001)
Food $_s$(1_year_before) $_{r,t}$*	-0.008	-0.010	-0.005	-0.007	-0.005	-0.007
	(0.015)	(0.015)	(0.014)	(0.014)	(0.015)	(0.015)
Industry controls	no	yes	no	yes	no	yes
F test on Food(global_chain) = Food *(1_year_before)*	14.90	10.62	n.a.	n.a.	n.a.	n.a.
p-value of F test	0.000	0.002	n.a.	n.a.	n.a.	n.a.
R-squared	0.019	0.02	0.02	0.021	0.019	0.02
No. of obs.	221236	220002	221236	220002	221236	220002

All models include a firm's age, firm fixed effects and region-year fixed effects. Industry-level controls include: $\ln(\text{imports})_{t-1}$, $\ln(\text{exports})_{t-1}$ and Herfindahl index $_t$.
Standard errors, clustered at the region-year level, are reported in parentheses.
* significant at 10%, ** at 5%, *** at 1%
n.a. stands for non applicable

Table 8 Strict Exogeneity Test
Fixed effects, TFP index

	chains present		**ln(no. of outlets)**		**ln(selling space)**	
Lag	0.023**	0.021**	0.020*	0.018*	0.003**	0.002**
$Food_s$*(global_chain) $_{r,t-1}$	(0.010)	(0.009)	(0.010)	(0.010)	(0.001)	(0.001)
Contemporaneous	0.018	0.017	0.013	0.014	0.002	0.002
$Food_s$*(global_chain) $_{r,t}$	(0.016)	(0.015)	(0.013)	(0.013)	(0.002)	(0.002)
Lead	-0.003	-0.006	-0.009	-0.01	-0.001	-0.001
$Food_s$*(global_chain) $_{r,t+1}$	(0.020)	(0.020)	(0.012)	(0.012)	(0.002)	(0.002)
Industry controls	no	yes	no	yes	no	yes
R-squared	0.020	0.021	0.02	0.021	0.02	0.021
No. of obs.	190620	189569	190620	189569	190620	189569

All models include a firm's age, firm fixed effects and region-year fixed effects. Industry-level controls include: $\ln(\text{imports})_{t-1}$, $\ln(\text{exports})_{t-1}$ and Herfindahl index $_t$.
Standard errors, clustered at the region-year level, are reported in parentheses.
* significant at 10%, ** at 5%, *** at 1%

Table 9 Controlling for Regional Wage
Fixed effects, TFP index

	chain present		ln(no. of outlets)		ln(selling space)	
$FOOD*(global\_chain)_{r,t-1}$	0.046***	0.036***	0.037***	0.034***	0.005***	0.004***
	(0.010)	(0.009)	(0.005)	(0.005)	(0.001)	(0.001)
$FOOD_s*(wage\_per\_L)_{r,t-1}$	-0.027	-0.033	-0.065	-0.065	-0.029	-0.033
	(0.053)	(0.050)	(0.052)	(0.051)	(0.052)	(0.050)
$ln(firm\ age)_{it}$	0.117***	0.117***	0.119***	0.118***	0.118***	0.117***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
$ln(imports)_{s,t-1}$		-0.029***		-0.019***		-0.027***
		(0.005)		(0.005)		(0.005)
$ln(exports)_{s,t-1}$		-0.003		-0.003		-0.003
		(0.004)		(0.004)		(0.004)
$Herfindahl\ Index_{st}$		-0.193***		-0.222***		-0.200***
		(0.041)		(0.040)		(0.041)
R-squared	0.019	0.02	0.02	0.021	0.019	0.02
No. of obs.	221236	220002	221236	220002	221236	220002

All models include firm fixed effects and region-year fixed effects.
Standard errors, clustered at the region-year level, are reported in parentheses.
* significant at 10%, ** at 5%, *** at 1%