Trade liberalization and inter-industry productivity spillovers: an analysis of the 1989-1998 Brazilian trade liberalization episode

Paz, Lourenco

Syracuse University

15 February 2012

Trade Liberalization and Inter-industry Productivity Spillovers:

An Analysis of the 1989-1998 Brazilian Trade Liberalization Episode

February 15, 2012

Lourenço S. Paz*
lspaz@maxwell.syr.edu
110 Eggers Hall
Syracuse University
Syracuse, NY 13244 USA

Abstract

The desire to increase manufacturing productivity has been a commonly cited goal of the trade liberalization episodes that have swept several developing countries since the 1980s. The literature has found evidence supporting such an increase in productivity. However, this paper finds that the methodology used in the literature ignores inter-industry productivity spillovers and suggests that this omission biases estimates of the impact of import tariff reduction on industry-level productivity. The findings from a case study of the Brazilian trade liberalization episode (1989-1998) indicate that the literature has overestimated the direct effect of trade liberalization by at least 25%, and that inter-industry productivity spillovers exist, are positive, and account for 70% of the increase in productivity that results from a reduction in import tariffs.

Keywords: productivity, trade liberalization, Brazil, spillovers

JEL Codes: F1, O3, L6

* I would like to thank Derek Laing, Badi Baltagi, Chihwa Kao, and Devashish Mitra for helpful discussions, Marc-Andreas Muendler for providing the Brazilian industry-level productivity series, Honório Kume for providing Brazilian import tariff data, and Wael Moussa for excellent research assistance. Thanks are also due to seminar participants at the Midwest International Trade Meeting. Any remaining errors are mine.

1. Introduction

The relationship between trade liberalization and productivity in developing countries has received considerable attention from researchers because productivity is a key factor for economic growth. Several studies have found that trade liberalization appears to increase industry-level productivity.[1] In addition, the literature indicates that trade liberalization improves firms' access not only to cheaper and better quality inputs, but also to a larger variety (e.g., Goldberg et al. 2010). In particular, a decrease in the input tariff has been found to increase industry-level productivity in Indonesia (Amiti and Konings (2007)), India (Khandelwal and Topalova (2011)), and Brazil (Schor (2004)).

However, the literature has failed to include the role of inter-industry productivity spillovers in analyses of the impact of tariff reductions on industry-level productivity. These spillovers may occur because of increased competition in the input market, which forces domestic input producers not only to increase their efficiency, but also to upgrade the quality of their products and even to embody more knowledge in the inputs produced, for instance by imitating the newly-imported competing intermediate inputs. Such improvements will be reflected in the input producers' productivity, which will spill over to the downstream industries through linkages that can be seen clearly in the Input-Output (I-O) matrix.[2]

There are at least two problems with the omission of productivity spillovers. The most obvious and serious problem is that the current methodology leads to inconsistent estimates of the effects of tariffs on productivity. If one of the main justifications for trade reform is to

[1] More specifically, several studies, including Krishna and Mitra (1998) for India, Pavcnik (2002) for Chile, Fernandes (2007) and Karacaovali (2011) for Colombia, and Ferreira and Rossi (2003) and Muendler (2004a) for Brazil, found that productivity increased as a result of a decline in output tariffs.
[2] The issue of productivity spillovers has been raised previously in the foreign direct investment literature (see Javorcik (2004)).

enhance productivity, it is very important to obtain a consistent measure of its impact. This inconsistent estimate is misleading for policymakers when they are considering a trade policy change for only a few industries, in which spillovers will not exist. The second problem concerns the interpretation of the estimated import tariff coefficient, since it is not clear what is being measured: the direct effect on productivity, the productivity spillovers received, or both.

This paper addresses an important gap in the literature by examining how inter-industry productivity spillovers affect estimates of the impact of trade liberalization on productivity. Using spatial econometric techniques, I assess the existence and magnitude of inter-industry productivity spillovers due to trade liberalization and the extent to which the omission of such industry interactions biases the estimated coefficient of the impact of tariffs on productivity. I also account for the endogeneity of tariffs with respect to productivity.[3]

To explore these issues, I use data from the Brazilian trade liberalization episode (1989-1998). The Brazilian experience offers a good benchmark because it has been studied extensively (see e.g., Muendler (2004a) and Schor (2004)) and the findings indicate that trade liberalization increased industry-level productivity. In particular, Muendler (2004a) estimated industry-level TFP using firm-level data, providing good quality estimates that I use in this paper.

The results of my analysis indicate that inter-industry spillovers are present, are positive, and can account for 70% of the increase in industry-level TFP due to trade liberalization. Contrary to the findings in the previous literature, my results indicate that when input tariffs are taken into account, their contribution to the increase in TFP due to trade liberalization is similar to the contribution of output tariffs. Furthermore, I find that not accounting for productivity

[3] This issue has generally been ignored in the literature. One exception is Karacaovali (2011), which accounts for the endogeneity of tariffs in an analysis of Colombia.

spillovers causes the impact of trade liberalization on productivity to be overestimated by at least 25%.

The remainder of the paper is organized as follows. The next section discusses in more detail the weaknesses of the methodology that has been used in previous studies and describes the methodology applied in this paper. The data set used in the empirical analysis is described in section 3. Section 4 reports my estimates and discusses the results. Finally, conclusions are presented in section 5.

2. Literature Review and Methodology

This section discusses the baseline specification that has been used in the literature and describes the methodology used in this paper, which accounts for inter-industry spillovers and addresses other problems with the baseline specification.

2.1 The Baseline Specification

Early studies (e.g., Pavcnik (2002) and Krishna and Mitra(1998)) focused specifically on the effect of output tariffs on industry-level productivity, since a decrease in output tariffs leads to increased competition in domestic markets and thus reduces x-inefficiencies through managerial restructuring. Industry-level productivity may also increase due to output reallocation from low productivity firms to high productivity firms, as shown in the Melitz (2003) model. The measure of productivity used in both these studies and the subsequent literature is total factor productivity (TFP), which measures the change in the level of output that cannot be explained by changes in the quantity of factors of production, i.e. capital, intermediate inputs (materials), and labor. This residual is composed of random shocks, process innovations, managerial effort and reorganization, increases in workers' knowledge, and knowledge embodied in intermediate inputs, which are all unobservable to the researcher

Later, the research agenda shifted toward examining the effect of input tariffs on productivity. There are two important reasons to focus on the role of input tariffs. First, Corden (1971) predicts that a decrease in input tariffs will decrease productivity because cheaper inputs will weaken competitive pressure in the output market (i.e., the effective rate of protection will increase). Second, the theoretical models in Ethier (1982), Markusen (1989), and Grossman and Helpman (1991) predict that lower input tariffs induce higher productivity through access to a larger variety and better quality of intermediate inputs as well as knowledge spillovers. These theoretical results suggest that the overall effect of the input tariff on TFP is an empirical question.

The input tariff effect is estimated in Schor (2004) for Brazil, Amiti and Konings (2007) for Indonesia, and Khendalwal and Topolova (2011) for India. Schor (2004) and Amiti and Konings (2007) define the input tariff as the weighted average of output tariffs computed using equation (1),

$$input\_tariff_{it} \equiv \sum_j \omega_{ji} * output\_tariff_{jt} \tag{1}$$

where the weight (ω_{ji}) is the share of input j in the total input cost of industry i's output and is assumed to be fixed during the period under study. Schor (2004) used weights from the I-O matrix, while Amiti and Konings (2007) obtained information on actual input purchases by firms, which allowed them to develop a more disaggregated measure of input tariffs than is possible using I-O matrices.[4] These three empirical studies found the coefficient to be negative, and that the effect of the input tariff on productivity was between two times and six times larger than the effect of the output tariff.

All the empirical studies discussed earlier assessed the relationship between trade liberalization and TFP in developing countries by estimating an econometric specification similar to equation (2), which, throughout this paper, is referred to as the baseline specification.

$$tfp_{it} = c + \gamma_1 * output\_tariff_{it} + \gamma_2 * input\_tariff_{it} + \boldsymbol{x}_{it}\beta + \Sigma_i\, \alpha_i + \Sigma_t\, \theta_t + u_{it} \qquad (2)$$

where tfp_{it} is the estimated industry i total factor productivity at time t, c is the constant term, $output\_tariff_{it}$ is the import tariff charged on the output produced by industry i at time t, $input\_tariff_{it}$ is the import tariff charged on the inputs used by industry i at time t, $\boldsymbol{x}_{it}$ is a vector of other control variables such as upstream TFP, α_i represents industry fixed effects, and θ_t represents year fixed effects.

Industry fixed effects are used in equation (2) for two reasons. First, TFP levels are expressed in conceptual units that are not comparable across industries. As a result, the identification has to come from within-industry variation, achieved by using either industry effects or estimating the first difference of equation (2). Second, industry effects account for some possibly omitted time invariant and industry-specific characteristics that are correlated with right-hand side variables, and thus prevent inconsistent estimates. One example of such industry-specific characteristics is labor or environmental regulations that affect industries differently and may constrain adjustment in some factors of production.

[4] Khandelwal and Topolova (2011) constructed the input tariff slightly different than Schor (2004). In lieu of the cost share, they used the share of the input in the value of the output.

Year effects are included in the specification to control for economy-wide shocks (i.e., variables that increase or decrease together in different industries during the same business cycle). For instance, if firms are prone to conduct managerial reorganizations during a recession, but at the same time the government raises tariffs in response to the recession, a spurious relationship will be found between tariffs and productivity unless year effects are used.

Fernandes (2007) argues that the estimated TFP series exhibits time persistence since the TFP estimation methodology in Olley and Pakes (1996) assumes that plant productivity follows a first-order Markov process. Unless the baseline specification accounts for this issue of time persistence, the estimates will be inconsistent, because the error term will be autocorrelated. Thus it is necessary to estimate an augmented version of the baseline model that incorporates the TFP of the previous period as a regressor and uses the Arellano-Bond estimator, where the estimated equation is in first-difference, as shown in equation (3), hereafter called the augmented baseline specification.[5]

$$\Delta tfp_{it} = \lambda \Delta tfp_{i,t-1} + \gamma_1 \Delta output\_tariff_{it} + \gamma_2 \Delta input\_tariff_{it} + \Delta \mathbf{x}_{it}\beta + \Sigma_t \, \Delta\theta_t + \Delta u_{it} \qquad (3)$$

The coefficient of the time-lagged TFP, λ, is expected to be positive because current TFP depends positively on the last-period TFP. Notice that the Δ is the time difference operator, for example, $\Delta tfp_{it} \equiv tfp_{it} - tfp_{it-1}$.

2.2 Shortcomings of the Baseline and Augmented Baseline Specifications

[5] This approach was used by Fernandes (2007) and Khandelwal and Topolova (2011).

There are at least two problems with the baseline specification and its augmented version. First, trade liberalization (i.e. the change in import tariffs) may not be exogenous with respect to TFP, as pointed out by Karacaovali (2011). Karacaovali (2011) developed a theoretical model in which industry TFP is an important political economy factor that affects tariff setting, reflecting the fact that the higher the current and expected future industry TFP, the greater the benefits of protection for firms in that industry, and thus the greater the incentives to lobby government for more protection. Using data from Colombia's trade liberalization episode, Karacaovali (2011) finds empirical evidence for the predictions in his theoretical model.[6] Hence, to obtain consistent estimates, the researcher needs to use an excluded instrument for tariffs.

The second problem is the failure to account for inter-industry spillovers, which leads to inconsistent estimates. To illustrate, I focus on the baseline specification, rewritten in vector notation, in which the variables in bold are vectors containing observations of all industries at year t, as shown in equation (4).[7]

$$\boldsymbol{tfp}_t = \boldsymbol{c} + \gamma_1 * \boldsymbol{output\_tariff}_t + \gamma_2 * \boldsymbol{input\_tariff}_t + \boldsymbol{x}_t \beta + \boldsymbol{\alpha}_t + \boldsymbol{\theta}_t + \delta \boldsymbol{Wtfp}_t + \boldsymbol{u}_t \qquad (4)$$

where the productivity spillovers are represented by the term $\boldsymbol{Wtfp}_t$. $\boldsymbol{W}$ is the matrix that contains the exogenous measure of influence (weights) of other industries' ($j \neq i$) TFP on industry *i's* TFP. The main diagonal of $\boldsymbol{W}$ contains only zeros because the within-industry spillovers are already included in the industry-level TFP measure.

Let $\boldsymbol{X}_t$ be the matrix of all regressors, i.e. $\boldsymbol{X}_t = (\boldsymbol{c}, \boldsymbol{output\_tariff}_t, \boldsymbol{input\_tariff}_t, \boldsymbol{x}_t, \boldsymbol{\alpha}_t, \boldsymbol{\theta}_t)$. Suppose the econometrician omits the spillover term. Then, the new error term will be given by

[6] Khandelwal and Topolova (2011) also found similar empirical evidence for the Indian tariff reform after 1998.
[7] This same reasoning applies to the augmented version of the baseline specification.

$\boldsymbol{\varepsilon}_t = \delta \boldsymbol{W} \boldsymbol{tfp}_t + \boldsymbol{u}_t$. Consistent estimates of the parameters require that $E[\boldsymbol{X}_t'\boldsymbol{\varepsilon}_t] = 0$. But the existence of spillovers prevents this condition from being met because $E[\boldsymbol{X}_t'\boldsymbol{\varepsilon}_t] = E[\boldsymbol{X}_t'\delta\boldsymbol{W}\boldsymbol{tfp}_t] + E[\boldsymbol{X}_t'\boldsymbol{u}_t] = E[\boldsymbol{X}_t'\delta\boldsymbol{W}\boldsymbol{tfp}_t] \neq 0$, since $\boldsymbol{tfp}_t$ is a function of $\boldsymbol{X}_t$. In other words, by not accounting for the spillover term, the regressors become correlated with the error term, leading to inconsistent estimates.

2.3 Addressing Endogeneity of Tariffs and Productivity Spillovers

To obtain consistent estimates of the effect of trade liberalization, I develop a methodology that addresses the two econometric problems discussed earlier: endogeneity of tariffs and TFP spillovers.

First, the reverse causality between TFP and tariffs has an important implication for the selection of instruments for import tariffs. The literature generally uses the pre-reform tariff level as an instrument for changes in tariffs (see e.g., Ferreira and Rossi (2003)). However, Karacaovali (2011) argues that the exclusion restriction for using the pre-reform tariff level as an instrument for tariff changes is not met, because the pre-reform tariff level takes into account the industry-level TFP present at that time, which in turn is correlated with the current TFP level. This means that the instrument (pre-reform tariff level) and the error term are correlated.

Interestingly, Muendler (2004a) found that one of the key goals of the Brazilian trade liberalization was to improve productivity in lagging industries. This suggests that a lower pre-reform productivity level implies a larger tariff cut. Thus, there is evidence that political economy factors may be driving trade liberalization in Brazil. Muendler (2004a) also found that TFP has some time persistence. So the argument in Karacaovali (2011) that the pre-reform tariff level cannot be used as an instrument for changes in tariffs also applies to Brazil.

To address this problem, I use Colombia's import tariffs during its trade liberalization episode (1984 to the mid-1990s) as an instrument for Brazil's import tariff.[8] Prior to their trade liberalization episodes, both the Colombian and Brazilian governments believed that their import substitution industrialization policies (which implied high levels of trade protection) were welfare enhancing, in addition to the fact that import substitution was viewed as an institution or even a historical legacy that could not be changed due to political concerns. At a certain point, however, governments realize that the gains from import substitution may not be as large as expected, and change their development policies by decreasing trade protection across all industries.[9] This means that the Brazilian and Colombian tariffs should exhibit a positive correlation, and thus move in the same direction (downward) as a result of this change to a trade liberalization policy.[10]

I believe that using Colombia's import tariffs as an instrument for Brazil's import tariffs is valid for the following reasons. First, Colombian tariffs are not affected by future Brazilian tariffs, since trade between these two countries is very small relative to their other partners. Second, the pre-reform protection patterns in Brazil and Colombia were different; Colombia offered more protection to the low productivity light manufacturing sector (e.g. apparel and footwear), while Brazil offered more protection to high productivity capital good industries (e.g. machinery and transportation equipment). Therefore, the possible effect of Colombia's productivity on its tariffs appears to be uncorrelated with the possible effect of Brazil's productivity on its tariffs.

[8] I match the year preceding the trade liberalization in Colombia (1984) to the year preceding trade reform in Brazil (1989). Hence, the 1984 Colombian tariff level is used as an instrument for the 1989 Brazilian import tariff, and so on.

[9] For example, governments that adopted import substitution development policies may have observed that countries with trade-oriented development policies, like South Korea, have experienced higher levels of economic growth.

[10] Karacaovali (2011) develops a similar argument in a theoretical model of the political economy of protection.

Second, the issue of TFP spillovers must be addressed. The existence of linkages across industries can be seen clearly in the I-O matrix through the amount of inputs that one industry purchases from other industries. This makes industry-level productivity dependent on the productivity of other industries, which works through two channels. The first channel is learning transfer (or technological diffusion), since technological knowledge can be embodied in intermediate goods.[11] The second channel is the upgrading of the quality of inputs purchased, due either to new requirements from buyers or imitation of newly-imported competing products.

The literature has controlled for such spillovers by adding a regressor that consists of the TFP of upstream industries aggregated according to their shares in the I-O table.[12] However, this approach does not provide consistent estimates because the industries are interdependent, so that upstream industry TFPs are simultaneously determined with the downstream industry TFP. To address this simultaneity problem, I use spatial econometrics techniques to explicitly model how much one industry interacts with another. Thus I include a spatial lag of the left-hand side variable in the augmented specification.

The amount of interaction between industries is reflected in the weights matrix $\boldsymbol{W}$, which is an $N \times N$ matrix with rows normalized to sum one and zeroes in the main diagonal.[13] The last requirement is needed to allow for identification in the estimated model. Intuitively, the elements of $\boldsymbol{W}$ should be larger for industries that have larger interactions. One way to measure this interaction is to use the share of inputs purchased by industry i from industry j (given by the I-O matrix) as the weights.[14] This new specification is shown in equation (5), where $\boldsymbol{tfp}_t$ is an N×1

---

[11] Keller (2009) provides a very good literature review on this topic and also discusses evidence that supports the argument that imports are an important channel of technology diffusion.

[12] For instance, Javorcik (2004) uses this approach, but the dependent variable is firm output level.

[13] This guarantees that the spatial lag coefficient (δ) will belong to the (-1,1) and allows it to be interpreted as the spatial multiplier as in Anselin (2003).

[14] The I-O matrix has been used previously in this way by Moreno et al (2004) among others.

vector of observations on the TFP (dependent variable), the spatial lag is given by $\boldsymbol{Wtfp}_t$, and δ is its estimated coefficient.

$$\Delta \boldsymbol{tfp}_t = \delta \boldsymbol{W} \Delta \boldsymbol{tfp}_t + \gamma_1 \Delta \boldsymbol{output\_tariff}_t + \gamma_2 \Delta \boldsymbol{input\_tariff}_t + \Delta \boldsymbol{\theta}_t + \Delta \boldsymbol{u}_t \tag{5}$$

Equation (5) uses the variables in first difference because, as discussed earlier, log TFP is not directly comparable across industries. However, the change in log TFP is comparable. Thus $\boldsymbol{W}\Delta\boldsymbol{tfp}_t$ can be interpreted as the term that captures the inter-industry spillover. I expect productivity spillovers to be positive, and thus predict that δ will also be positive.

One important issue concerns the potential endogeneity of $\boldsymbol{W}$. I address this issue by using the 1985 I-O matrix, which is five years before the start of the trade liberalization. It is worth mentioning that the 1985 I-O matrix is very similar to the 1990 and 1995 I-O matrices. This means that if there is indeed endogeneity, it appears to be time invariant. In particular, as discussed earlier, the assumption that the input mix is stable over time seems reasonable. Thus, in this case, industry effects or using the variables in first difference solves the problem.

Given that the spatial lag is an endogenous regressor, I use the Generalized Spatial Two Stage Least Square (GS2SLS) estimator, and follow the suggestion in Kelejian and Prucha (1998) to use $\boldsymbol{W}_N\boldsymbol{h}_{it}$ and $\boldsymbol{W}^2_N\boldsymbol{h}_{it}$ as instruments for the spatial lag, where $\boldsymbol{h}_{it}$ is a vector of exogenous regressors (included instruments) and excluded instruments (Colombian tariffs).[15] The final specification combines both the time and spatial lags, as shown in equation (6).

$$\Delta \boldsymbol{tfp}_t = \lambda \Delta \boldsymbol{tfp}_{t-1} + \delta \boldsymbol{W}_N \Delta \boldsymbol{tfp}_t + \gamma_1 \Delta \boldsymbol{output\_tariff}_t + \gamma_2 \Delta \boldsymbol{input\_tariff}_t + \Delta \boldsymbol{\theta}_t + \Delta \boldsymbol{u}_t \tag{6}$$

[15] More specifically, I used the Stata command spivreg, which is described in detail in Drukker et al. (2011).

As a robustness check, I assume that Δu_{it} presents spatial correlation of the following form: $(1-\rho \mathbf{W}_N)^{-1}\Delta \boldsymbol{u}_{it} = \Delta \boldsymbol{e}_{it}$, where $\boldsymbol{e}$ is a vector of i.i.d. error term with zero mean and finite variance. I estimate equation (6) by generalized method of moments (GMM), similar to a spatial Arellano-Bond model, in which the instruments for the lagged TFP are the Arellano-Bond instruments and the instruments for the spatial lag of TFP are the same as those suggested by Kelejian and Prucha (1998). According to Elhorst (2010), the advantage of using GMM is that it imposes a significantly smaller computational burden and prevents serious numerical problems when finding the eigenvalues of the weights matrix, which is required in the maximum likelihood estimator.

3. Policy Background and Data Description

This section provides some historical background on the 1989-1998 Brazilian trade liberalization and presents the data used in the econometric analysis of this episode. First, I explain how Brazilian trade policy changed in the 1990s. This is followed by a detailed description of the TFP data. Then, I describe the I-O table, the weighting matrices (W), and the upstream TFP variables used in the estimates. Finally, I discuss the sources of the Brazilian and Colombian import tariff data.

3.1 The 1989-1998 Brazilian Trade Liberalization

Until the end of the 1980s, Brazil's trade policy was dictated by an import substitution development policy and the country's balance of payments deficits.[16] The former implied different levels and types of protection across industries, in particular high tariffs and non-tariff barriers (NTBs) on imported goods that competed with similar domestic products. The latter resulted in increased tariffs across all industries in order to curb imports and generate trade balance surpluses. Moreover, Brazil used its developing country status under article XVIII of the General Agreement on Tariffs and Trade (GATT) to avoid participating in all tariff reduction rounds.

This trade policy started to change in 1988 when Brazil unilaterally decided to decrease the level of redundant protection. Tariffs were reduced to a level that would still curb imports, but, as stressed by Kume et al. (2003), no NTBs were eliminated. In 1990, Brazil's new president drastically reduced NTBs and adopted nominal tariff reductions scheduled to start in 1990 and end in 1994. The actual decrease in tariffs was not identical across industries. Moreover, the tariff reductions did not follow the planned schedule. Nonetheless, the tariff reductions had real effects on the economy, as imports of manufactured goods increased by more than 200% and import penetration increased from 5.7% to 11.6% between 1990 and 1998.

3.2 TFP Data

The industry-level TFP series are from Muendler (2004a), who used firm-level data from the Pesquisa Industrial Annual, an annual survey conducted by the Instituto Brasileiro de Geografia e Estatística (IBGE) that consists of an unbalanced panel of roughly 9500 medium-

[16] Kume et al. (2003) provide a comprehensive description of Brazil's trade policy in the 1980s and 1990s.

and large-sized firms.[17] Muendler (2004a) estimated the TFP for the 1986-1998 period using two methods.[18] The first TFP measure (OLS-TFP) is the estimated OLS residual ($\widehat{\varepsilon_{it}}$) of a simplified production function, shown in equation (7).

$$y_{it} = \beta_{bl}^{ols} l_{it}^{bl} + \beta_{wh}^{ols} l_{it}^{wh} + \beta_{K}^{ols} k_{it} + \beta_{S}^{ols} s_{it} + \beta_{M}^{ols} m_{it} + \varepsilon_{it} \tag{7}$$

where all variables are expressed in natural logarithms, and y is the output; l^{bl} is the number of blue-collar workers; l^{wh} is the number of white-collar workers; k is the stock of equipment; s is the stock of structures that encompasses real state, premises, vehicles, computers, and rented or leased capital goods; m is the amount of materials (intermediate inputs) used in production; and ε is the error term.

The second TFP measure (OP-TFP) in Muendler (2004a) is estimated using an extended version of the Olley and Pakes (1996) methodology that takes into account the investment decision concerning the two types of capital (equipment and structures) and endogenous firm entry and exit. The OP-TFP is depicted by equation (8).

$$y_{it} = \beta_{bl} l_{it}^{bl} + \beta_{wh} l_{it}^{wh} + \beta_{K} k_{it} + \beta_{s} s_{it} + \beta_{M} m_{it} + \omega_{it} + \varepsilon_{it} \tag{8}$$

where ω is the time-varying firm's TFP (unobservable). The extended Olley-Pakes strategy used by Muendler (2004a) produces consistent estimates of TFP by addressing the well-known problem of the correlation between ω and the input factors.[19]

---

[17] The survey was not conducted in 1991 due to budget cuts, so I used linear interpolation to build the TFP for 1991. More details about the survey and its variables can be found in Muendler (2003).
[18] The interested reader can refer to Van Beveren (2011) for a survey on TFP estimation.

Muendler (2004a) aggregated the firm-level TFP data at IBGE's nível 50 industry classification (27 manufacturing industries). I further aggregated the data by simple average to 16 manufacturing industries because the nível 50 classification is incompatible with the ISIC classification used for other variables (discussed later).

The correlation and descriptive statistics concerning the two TFP measures are presented in Tables 1 and 2, respectively. The two measures are calculated using the unbalanced panel sample of all firms (All) and the balanced panel sample containing only those firms that stayed in the market, i.e. had positive output, during the whole trade liberalization period (Stayers). Even though the coefficients of the production inputs vary across the different measures, the TFP estimates exhibit very similar behavior, as indicated by the high correlation among the measures (Table 1) and the descriptive statistics (Table 2).

3.2 I-O Table, Weighting Matrix, and Upstream TFP Variables

The I-O table is used to calculate the weighting matrix, the upstream TFP variable, and the input tariff, with the input tariff calculated according to equation (1). I use the 1985 I-O table for Brazil from IBGE (2006). The non-manufacturing sectors and all final use consumption columns are excluded. The original table used nível 80 industry classification (48 manufacturing industries). However, I further aggregated it to 16 manufacturing industries in order to match the industry aggregation level dictated by the tariff data. This procedure leads to an I-O matrix (Γ) of dimensions 16x16.

The weighting matrix $\boldsymbol{W}$ has zeroes in its main diagonal. The other entries (w_{ij}) are the share of inputs purchased from industry j by industry i for $i \neq j$, as displayed by equation (9).

---

[19] A brief description of this strategy is presented in the appendix. A detailed explanation of the estimation procedure and its theoretical derivation can be found in Muendler (2004b).

$$w_{ij} = \frac{\Gamma(i,j)}{\sum_{k=1}^{16}\Gamma(k,j)}, with\ \ i\ ,j\ =\ 1,\ldots,16 \qquad (9)$$

Following Karacaovali (2011), the upstream TFP variable is calculated using the share of each upstream industry in the I-O matrix multiplied by the respective upstream industry TFP.

3.3 Tariff Data

The 1987-1998 Brazilian import tariff data set is from Kume et al. (2003) and was originally aggregated from individual product tariff lines to IBGE's nível 50 industry classification using industry value-added as weights. Using the same type of weights, I further aggregated the Brazilian output tariff data into 16 manufacturing industries. As discussed earlier, the Brazilian input import tariff was constructed according to equation (1).

The instrument for Brazilian import tariffs was constructed using Colombian import tariff data for the 1982-1993 period (from the Colombian National Planning Department) aggregated at the 4-digit ISIC level. I further aggregated the data to my 16-industry classification using simple averages. Similarly, the instrument for the Brazilian input tariffs is calculated using equation (1), but now the output tariffs are those from Colombia.

4. Empirical Results

In this section, I first present estimates from the baseline specification concerning the effects of output and input tariffs on industry-level productivity. Next, I present the results when the time persistence in TFP is taken into account by the augmented baseline specification. I then estimate two specifications that address productivity spillovers across industries and provide consistent estimates of the effect of tariffs on industry-level productivity. All of these results were estimated using both TFP measures and the unbalanced panel sample (All). My results confirm the existence of (large) TFP spillovers. Moreover, these consistent estimates indicate that the effect of output tariffs on TFP is smaller relative to the specifications that ignore the existence of spillovers. The section concludes with some robustness checks that support these results, including estimates using the balanced panel sample (Stayers).

4.1 Estimates from Baseline Specifications

The results from the first baseline specification estimation by OLS are presented in Table 3. Columns (1)-(3) contain the estimates for the OLS-TFP measure and columns (4)-(6) contain the estimates for the OP-TFP measure. Columns (1) and (4) do not include year and industry fixed effects, and their output tariff coefficients are positive but statistically significant at the 10% level only for the OLS-TFP measure in column (1).

When year and industry effects are included in columns (2), (3), (5), and (6), the output tariff coefficients become negative, are always statistically significant at the 5% level, and have a point estimate of about -0.11. This change in signs occurs because time-invariant industry-specific characteristics matter for import tariff setting. The result is also consistent with the findings of Ferreira and Rossi (2003) for Brazil and Karacaovali (2011) for Colombia. Neither the upstream TFP coefficients nor the input tariff coefficients are statistically significant in any of the specifications. The result for the input tariff coefficients is the opposite of the negative and statistically significant coefficients for the input tariff found by Schor (2004) for Brazil using a different time period and TFP measure. The estimated coefficient of output tariff in column (2) means that a 10 percentage point decrease leads to a 1.13 percent increase in TFP, which is similar to the results in Schor (2004) and slightly larger than the results in Ferreira and Rossi (2003).

To address the potential simultaneity between import tariff and TFP, I estimated equation (2) with year and industry effects using instrumental variables with Colombian tariffs and the input tariff calculated using Colombian tariffs as excluded instruments; the latter is used whenever the Brazilian input tariff is an endogenous regressor. The results are shown in Table 4. All specifications have a negative and statistically significant coefficient for output tariffs. These coefficients are at least 75% larger than the OLS coefficients in Table 3. Although the coefficients of upstream TFP in Table 4 are positive, they are not statistically significant at the 5% level in any of the specifications. The coefficients of input tariff are positive and statistically significant at the 5% level, which is contrary to the findings in the literature.

In the first stage regressions of columns (1) and (4) in Table 4, Colombian tariff is always positive as expected and statistically significant at the 5% level. Upstream TFP is positive but statistically significant at the 5% level only when year and industry effects are not included. The Colombian tariff cannot be considered a weak instrument in columns (1) and (4) because the Kleibergen-Paap rK Wald F-statistic is at least 28, which is much larger than the Stock-Yogo 10% maximum IV relative bias critical value of 16.38. This is also the case for the remaining columns, which have a Kleibergen-Paap statistic of at least 12 and a Stock-Yogo critical value of 7. Exogeneity of the output tariff is rejected at the 5% level by the Hausman test for the OLS-TFP specification in column (1). For the OP-TFP specification in column (4), the p-value is 16.5%. For columns (2), (3), (5), and (6) the null hypothesis of the endogeneity test is that both output and input tariffs are exogenous, which is rejected at the 5% level for all specifications. These results support the approach of treating import tariff as an endogenous regressor.

Given that the estimates from equations (3), (5), and (6) are expressed in their first difference, for the sake of comparison, Table 5 reports the estimates of the baseline specification, equation (2), in first difference with year effects using the IV estimator. The output tariff coefficients are always negative but statistically significant at the 5% level only in columns (1)-(4) and at the 10% level in columns (5) and (6). The upstream TFP coefficients are always positive and not statistically significant at the 10% level. The input tariff coefficients are negative as expected but not statistically significant. The null hypothesis that both output and input tariffs are zero is rejected at the 5% level in all specifications.

In the first stage regression for columns (1) and (4), the Colombian tariff coefficients in Table 5 are positive and slightly larger in magnitude than the coefficients in Table 4, and statistically significant at the 5% level. The upstream TFP is positive and statistically significant at the 10% level. For all specifications, the Kleibergen-Paap weak identification Wald *F*-statistic is larger than the Stock-Yogo reference values, so the possibility of a weak instrument is not a concern here. The null hypothesis of exogeneity of output tariffs is rejected only at the 10% level for OLS-TFP, although the *p*-values for columns (2)-(4) are close to 20%.

4.2 Accounting for Time Persistence in TFP

The concern that TFP is likely to be autocorrelated in the Brazilian data is addressed by estimating equation (3) with an Arellano-Bond estimator, which uses lags of the endogenous regressors as instruments (hereafter called Arellano-Bond instruments) as well as Colombian tariffs.[20] The results are reported in Table 6.

The time-lagged TFP coefficient is positive and statistically significant only in columns (1) and (4), when the Arellano-Bond instruments consist of the 2nd, 3rd, and 4th lags of the endogenous regressor (tfp_{it-1}), in which the import tariff coefficients are not statistically significant. On the other hand, the time-lagged TFP is not statistically significant when the Arellano-Bond instruments use the 3rd and the 4th lags of the endogenous regressors (columns (2), (3), (5), and (6)). In this case, the output tariff coefficient is always negative and statistically significant. The upstream and input tariff coefficients are positive and not statistically significant at the 10% level. The null hypothesis that both output and input tariffs are equal to zero is not rejected at the 5% level of confidence for columns (3) and (6). The output tariff coefficient is larger than the IV coefficient for the specification in first difference when input tariff is not a regressor and smaller than the IV coefficient when input tariffs is a regressor.

Some of the lags used for the Arellano-Bond instruments may be invalid instruments if the error term still displays autocorrelation. The first type of test to detect this problem is a specification test known as an over-identification test. The null hypothesis of the over-identification test is not rejected at the 5% level only for column (5). However, this is not sufficient evidence to conclude that all the estimates are invalid because in small samples this test has the tendency to over-reject the null hypothesis. In addition, this test result does not offer any insight concerning the problematic lag(s).

Another type of specification test checks directly for the presence of autocorrelation in the estimated residual, and the order of the autocorrelation in the null hypothesis can be chosen by the researcher. Interestingly, the null hypothesis that the error term is not an AR(1) process is

[20] I used the Stata command xtabond2, described in detail by Roodman (2006).

rejected whenever the second lag of the endogenous regressors is used as an excluded instrument, whereas the null hypothesis that the error term is not an AR(2) process is never rejected. In particular, the rejection of the null hypothesis of the AR(1) implies that the second lag of the endogenous regressor is correlated with the residual and thus is an invalid instrument. This means that only instruments using the third lag and beyond would be valid. This supports the conclusion that the estimates that should be seriously considered are those that use the third and fourth lags, which provide no evidence of time persistence for the Brazilian TFP.

4.3 Accounting for Productivity Spillovers

Table 7 presents the results for equation (5), which accounts for the spatial interaction across industries (inter-industry spillovers) by using the GS2SLS estimator, and should therefore produce consistent estimates if the error term does not present autocorrelation over time. The TFP spatial lag is positive as expected and statistically significant at the 5% level in all specifications. The output tariff coefficient is negative and statistically significant at the 5% level in all columns except column (6), where it is significant at only the 10% level.

In columns (2), (3), (5), and (6) the error term is allowed to have a spatial correlation structure given by the weighting matrix. The correlation coefficients are negative and not statistically significant. The estimated effect of output tariffs on productivity is 10% to 40% smaller than the estimated effect in Table 5. The input tariff coefficient is positive in column (3) for OLS-TFP and negative for OP-TFP in column (6), with neither being statistically significant. The null hypothesis that both output and input tariffs are zero is rejected at the 5% level for both columns.

The results in Table 7 suggest that a 10% decrease in all tariffs would lead to an increase in TFP of 1.22% (based on column (4) estimates) through the direct effect of tariffs and an increase of 2.93% through the inter-industry spillover effect, which is calculated using the spatial multiplier defined in Anselin (2003): $(1\text{-}0.706)^{-1}$ -1 = 2.4. The spillover effect accounts for 70% of the total effect of tariffs on TFP. Although the spillover effect here seems very large, it is equivalent to only about half the effect of input tariffs on TFP (which is six times larger than the effect of output tariffs) found by Khandelwal and Topolova (2011).

Finally, Table 8 accounts for both TFP spillovers across industries and time dependence by estimating equation (5) using a Spatial Arellano-Bond estimator (cf. Elhorst et al. 2010). I also used the first spatial lag of Colombian tariffs as an instrument.[21] As shown in Table 8, the TFP time lag is positive and statistically significant at the 5% level in columns (1) and (3), as is the case in Table 6. The spatial lag TFP is always positive and statistically significant.

The output tariff coefficient is always negative, but in the OLS-TFP specifications it is significant only in column (2) and at the 10% level in columns (1) and (3). In the OP-TFP specifications it is statistically significant at the 5% level only in columns (4) and (5) and at the 10% level in column (6). In terms of magnitude and significance, the output tariff coefficients and the spatial lag of the TFP coefficients differ by less than one standard deviation from the figures in Table 7, where the input tariff coefficients are negative but not statistically significant. As is the case in Tables 6 and 7, whenever input tariffs is included in the estimated specification for the OP-TFP, the output tariff coefficient decreases in absolute size, as also found by Amiti and Konings (2007), and its standard error increases. The null hypothesis that both output and input tariffs are zero cannot be rejected at the 10% level for column 3, but is rejected at the 10%

[21] The second spatial lag of Colombian tariffs was not used due to collinearity in the instruments matrix.

level for the specification in column (6), which suggests that they both matter in determining TFP. However, they cannot be precisely estimated.

The null hypothesis of the overidentification test is rejected at the 5% level for column (1) and at the 10% level for columns (2) and (4). Furthermore, similar to the results in Table 6, the null hypothesis of the AR(1) test is rejected at the 5% level for the specifications using the 2^{nd}, 3^{rd}, and 4^{th} lags of variables as instruments, which are reported in columns (1) and (4). The null of the AR(2) test is never rejected. The results of the AR tests imply that only instruments based on the 3^{rd} and 4^{th} lags of regressors could be considered valid. In the end, column (3) for the OLS-TFP measure and columns (5) and (6) for the OP-TFP measure are the only estimates that are not rejected in the diagnostic tests. Most importantly, the spatial lag of TFP is positive and statistically significant and the time lag of TFP is not statistically significant in these three specifications.

4.4 Robustness Checks

I conducted three robustness checks of the empirical results above. First, I estimated all the previous regressions using the TFP measures obtained by considering the balanced panel sample, that is, only those firms that remained in the sample throughout the trade liberalization period (Stayers). The results were very similar, and are available upon request.

The second robustness check was to re-estimate the specifications in Table 7 using a heteroskedastic robust estimator for the coefficients' standard errors, as developed in Kelejian and Prucha (2007). The results of this exercise are reported in Table 9. With the exception of some very small changes in the magnitude of the estimated standard errors, the results do not change for the specifications without spatial correlation in the error term (columns (1) and (4)).

However, the specifications that account for spatial correlation in the error term (columns (2), (3), (5) and (6)) exhibit a slightly larger coefficient for the spatial lag of TFP and the error spatial correlation coefficient is again negative, but it is now statistically significant at the 5% level. The key result from Table 9 is that the spatial lag of TFP remains positive and statistically significant, as in Table 7.

The third robustness check addresses two issues raised in the literature (see Schor (2004) and Amiti and Konings (2007)). The first is the well-known problem of distinguishing between physical productivity and mark-ups in imperfectly competitive industries. The second is that the effect of tariff changes on TFP depends on the initial level of competitiveness present in the industry. This relies upon the assumption that the less competitive the industry is initially, the larger are the x-inefficiencies, and therefore the larger the possible TFP gain due to trade liberalization. To address these issues, Schor (2004) and Amiti and Konings (2007) added a measure of industry competitiveness (the Herfindahl index) to the baseline specification. But this approach has a serious drawback, because in a Melitz (2003) theoretical framework, industry-level productivity and market concentration are simultaneously determined. This means that a time and industry-varying Herfindahl index would be an endogenous regressor, a point that has been overlooked in the literature.[22]

Unfortunately, for Brazil, neither a good instrument for industry concentration nor the firm level information needed to compute the Herfindahl index for each industry-year pair is available. However, Schor (2004) reports the Herfindahl index for each industry in 1986 (i.e., three years before the trade liberalization). This allows me to address the issue of initial industry competitiveness by creating a dummy variable, hereafter called the Herfindahl dummy, to

[22] Moreover, as pointed out by Trefler (2003) for NTBs and Karacaovali (2011) for tariffs, industry concentration is an important factor in determining tariffs. So, it is not clear that the inclusion of a concentration measure will capture the desired effect.

distinguish industries with an initial low level of competitiveness from the others. The Herfindahl dummy is "1" for six industries (the top third) with the highest Herfindahl index (i.e., industries with highly concentrated markets) and "0" otherwise. Then, I re-estimated the baseline specification as in Table 5 using an interaction between the Herfindahl dummy and the output tariff. As discussed earlier, this coefficient is expected to be negative. The results are reported in Table 10.

Columns (1) and (4) present the OLS estimates and the remaining columns present the IV estimates, where the endogenous regressors are the output tariff, input tariff, and the interaction between the concentration dummy and output tariffs, using the interaction of the Herfindahl dummy and Colombian tariffs as well as the Colombian tariffs and the input tariff constructed using Colombian tariffs as excluded instruments. As in Table 5, the output tariff coefficients are negative and statistically significant at the 10% level and, in some cases, at the 5% level. Interestingly, the OLS coefficients (columns (1) and (4)) decreased by 50% in magnitude, while the IV coefficients in the remaining columns exhibited similar magnitudes to those in Table 5. The input tariff coefficients are not statistically significant. The coefficient of the interaction between the Herfindahl dummy and output tariff is negative but never statistically significant, even in the specifications estimated by OLS. The weak identification Kleibergen-Paap rK Wald statistic is smaller than the Stock Yogo reference values (see columns (2) and (5)). Although this indicates that the instruments used are weak, unfortunately, better instruments are not available.

Finally, I added the Herfindahl dummy interactions to the Table 9 specifications that take into account TFP spillovers. The results are presented in Table 11. The TFP spatial lag coefficient is always positive and statistically significant at the 5% level, and its magnitude is similar to the figures in Table 9. In Table 11 the output tariff, the interaction between the

Herfindahl dummy and the output tariff coefficient, and the input tariff are all negative (as expected), but not statistically significant in any of the specifications. Interestingly, the magnitude of the output tariff decreased and its standard deviation increased. This result was expected and is due to the fact that the variables are correlated and that, unfortunately, the number of observations is too small to estimate the coefficients precisely. In order to determine whether multicollinearity could, in fact, be a reason for the lack of precision, I conducted two tests of the specifications in columns (2) and (4). First, I tested the null hypothesis that the output tariff and its interaction with the Herfindahl dummy are jointly equal to zero. It is rejected at the 5% level for column (2) and at the 10% level for column (4), with a *p*-value of 5.6%. Second, I tested the null hypothesis that the output tariff, its interaction with the Herfindahl dummy, and the input tariff are jointly equal to zero. This hypothesis is also rejected at the 5% level for column (2) and at the 10% level for column (4), with a *p*-value of 7%. These results suggest that the tariffs and the output tariff interaction with the Herfindahl dummy jointly affect TFP. However, precise estimates cannot be obtained.

5. Summary and Conclusions

Several developing countries have used trade liberalization to boost productivity, and there has been a significant amount of research evaluating the relationship between trade liberalization and productivity. In several cases, the findings indicate that reductions in both output and input tariffs caused an increase in industry-level productivity (see e.g., Schor (2004), Amiti and Konings (2007), and Khandelwal and Topalova (2011)). However, these studies ignore the existence of inter-industry spillovers, which leads to inconsistent estimates. Moreover, this omission means that the results offer little guidance to policy-makers interested in narrow-scope trade liberalization or even a fine-tuning of trade policy, in which spillovers are very small or non-existent.

This paper has addressed this shortcoming of the literature by accounting for inter-industry spillovers. The methodology proposed here incorporates these inter-industry spillovers into the estimated model through spatial econometric techniques and also addresses the endogeneity of tariffs with respect to productivity. Applying this methodology to data for the Brazilian trade liberalization episode (1989-1998), I find that inter-industry spillovers not only exist but also have a positive effect on industry-level productivity. These spillovers can account for 70% of the increase in industry-level TFP due to trade liberalization. When the spillovers are ignored, the estimates of the impact of trade liberalization on productivity are biased upward by at least 25%. These results suggest that when designing trade policies that deal with only a few industries, policymakers should treat estimates that ignore spillovers with caution, since the actual increase in productivity may be considerably smaller.

References

Amiti, M. and J. Konings, (2007). Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia. American Economic Review, 97(5), 1611-1638.

Anselin, L. (2003). Spatial Externalities, Spatial Multipliers, and Spatial Econometrics. International Regional Science Review, 26 issue 2, 153-166.

Corden, Max W. (1971). The Theory of Protection. Oxford: Oxford University Press

Drukker, D. M., I. R. Prucha, and R. Raciborski. 2011. A command for estimating spatial-autoregressive models with spatial autoregressive disturbances and additional endogenous variables. Working paper, University of Maryland, Department of Economics.

Elhorst, J. P., (2010). Dynamic panels with endogenous interaction effects when T is small. Regional Science and Urban Economics, 40, 272-282.

Ethier, W., (1982). National and International Returns to Scale in the Modern Theory of International Trade. American Economic Review, 72(3), 389-405.

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

Ferreira, P. C. and J. L. Rossi (2003). New Evidence from Brazil on Trade Liberalization and Productivity Growth. International Economic Review, 44(4), 1383-1405.

Goldberg, P., A. Khandelwal, N. Pavcnik and P. Topalova. (2010). Imported Intermediate Inputs and Domestic Product Growth: Evidence from India. Quarterly Journal of Economics, 125(4), 1727-67.

Grossman, G. and E. Helpman. (1991). Innovation and Growth in the Global Economy. Cambridge, MA: MIT Press.

IBGE (2006). Notas Técnicas Matriz de Insumo-Produto Brasil – 1985. Instituto Brasileiro de Geografia e Estatística.

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), 605-27.

Karacaovali, B., (2011). Productivity Matters for Trade Policy: Theory and Evidence. *International Economic Review*, 52(1), 33-62.

Kelejian, H. and I. Prucha (1998). A generalized spatial two stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. Journal of Real Estate Finance and Economics, 17, 99–121.

Kelejian, H. and I. Prucha (2007). HAC Estimation in a Spatial Framework. Journal of Econometrics, 140, 131-154.

Kelejian, H. H., and D. P. Robinson (1993). A Suggested Method of Estimation for Spatial Interdependent Models with Autocorrelated Errors, and an Application to a County Expenditure Model," Papers in Regional Science, 72, 297-312.

Keller, W. (2009). International Trade, Foreign Direct Investment, and Technology Spillovers. NBER Working Paper No. 15442.

Khandelwal, A. and P. Topalova. (2011). Trade Liberalization and Firm Productivity: The Case of India. The Review of Economics and Statistics, 93(3), 995-1009.

Krishna, P., and D. Mitra, (1998). Trade liberalization, market discipline and productivity growth: new evidence from India. Journal of Development Economics, 56, 447-462.

Kume, H., G. Piani, and C.F. Souza (2003). Instrumentos de Politica Comercial no Periodo 1987-1998 in Corseuil, C. H. and H. Kume (eds.) "A abertura comercial brasileira nos anos 1990: impactos sobre emprego e salário". Rio de Janeiro: Ipea.

Markusen, J. (1989). Trade in Producer Services and in Other Specialized Intermediate Inputs. American Economic Review, 79(1), 85-95.

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

Moreno, R., López-Bazo, E., Vayá, E. And Artís, M. (2004). External Effects and Cost of Production. In Anselin, L., Florax, R. and Rey, S. (eds.) “Advances in Spatial Econometrics”. Berlin ; New York : Springer-Verlag.

Muendler, Marc-Andreas, (2003). The Database Pesquisa Industrial Anual 1986-2001: A Detective's Report. Instituto Brasileiro de Geografia e Estatística, Rio de Janeiro, Mimeograph (econ.ucsd.edu/muendler/brazil).

Muendler, Marc-Andreas, (2004a). Trade, Technology, and Productivity: A Study of Brazilian Manufacturers, 1986-1998. CESifo Working Paper Series 1148, CESifo Group Munich.

Muendler, Marc-Andreas, (2004b). Estimating Production Functions When Efficiency Choice Is Endogenous. University of California, San Diego, Mimeograph.

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

Pavcnik, N. (2002). Trade Liberalization, Exit, and Productivity Improvement: Evidence from Chilean Plants. Review of Economic Studies, Wiley Blackwell, vol. 69(1), 245-76.

Roodman, D. (2006). How to Do xtabond2: An introduction to "Difference" and "System" GMM in Stata. Working Paper 103, Center for Global Development, Washington.

Schor, A. (2004). Heterogeneous Productivity Response to Tariff Reduction: Evidence from Brazilian Manufacturing Firms. Journal of Development Economics, 75(2), 373-96.

Sethupathy, G. (2007). Does Exporting Lead to Productivity Spillovers in Horizontal or Vertical Industries? Evidence from Indonesia. Discussion Paper No.: 0708-01. Department of Economics, Columbia University.

Van Beveren, I. (2011). Total Factor Productivity Estimation: A Practical Review. Journal of Economic Surveys, 26(1), 98-128.

Appendix: Extended Olley-Pakes Methodology

The Olley-Pakes methodology was developed to address two well-known endogeneity problems that arise when estimating production functions. The first problem is the correlation between the unobservable productivity shocks (ω_{it}) and the quantities of inputs that are chosen by the firm.[23] When ignored, this correlation leads to inconsistent estimates, as, for instance, if OLS is used to estimate equation (7). The second problem is sample selection due to productivity levels, which occurs because firms that exit the market do so when their productivity (ω_{it}) falls below a certain level. As a result, the surviving firms' ω_{it} will come from a selected sample, which affects the level of inputs used. Muendler (2004a,b) extended the Olley-Pakes methodology to account for two types of investment in capital: equipment and structures.[24] Thus, the production function to be estimated is equation (8):

$$y_{it} = \beta_{bl} l_{it}^{bl} + \beta_{wh} l_{it}^{wh} + \beta_K k_{it} + \beta_s s_{it} + \beta_M m_{it} + \omega_{it} + \varepsilon_{it} \qquad (8)$$

where y_{it} is the log of output; l^{bl} is the log of the number of blue-collar workers; l^{wh} is the log of the number of white-collar workers; k_{it} is the log of total equipment; s is the stock of structures that encompasses real state, premises, vehicles, computers, and rented or leased capital goods; m_{it} is the log of total materials; and ε is the error term.

Following the Olley-Pakes methodology, let both capital and structures be accumulated by firms through a deterministic dynamic investment process (I^K and I^S, respectively) that arises

[23] Productivity shocks are assumed to be under the control of the firm's management but are unobservable by the researcher.

[24] Muendler (2004a) also tried to account for potential efficiency difference between foreign and domestic equipment and intermediate good inputs that would otherwise be attributed to overall TFP. But his findings indicate there is no such difference.

from the firm's profit maximization problem.[25] These investment functions depend on time *t* state variables such as the current stock of capital (equipment or structures), current productivity. and variables representing not only the economic environment, but also the firms' individual expectations about market demand. These latter variables ($\mathbf{D}_t$), which characterize a firm's environment, are foreign market penetration, the economy-wide real exchange rate, nominal tariffs, aggregate demand, and the annual inflation rate. To prevent a simultaneity problem from changes in D_t due to changes in productivity, Muendler (2004a) uses the nominal exchange rate and foreign producer price indices at the sector level as instrumental variables to predict foreign market penetration and nominal tariffs. Since the investment functions are invertible, let ω_{it} be described by equation (10), where both β_{0i} and ξ_{it} are known to the firm when it chooses variable factor inputs and investment for the next period.

$$\omega_{it} = h\left(I_{it}^{k}, I_{it}^{s}, a_{it}, k_{it}, s_{it}, k_{it}^{F}; \boldsymbol{D}_{it}\right) + \beta_{0i} + \xi_{it} \tag{10}$$

Once ω_{it} is fully characterized, the TFP estimation strategy uses three regressions. The first regression equation is given by equation (11).

$$y_{it} = \beta_{0i} + \beta_{bl} l_{it}^{bl} + \beta_{wh} l_{it}^{wh} + \beta_M m_{it} + \phi\left(I_{it}^{k}, I_{it}^{s}, a_{it}, k_{it}, s_{it}, k_{it}^{F}; \boldsymbol{D}_{it}\right) + \xi_{it} + \varepsilon_{it} \tag{11}$$

where a polynomial series estimator of fourth-order approximates the following function: $\phi(.) \equiv \beta_K k_{it} + \beta_s s_{it} + h(.)$. Each firm's individual productivity is estimated, which provides

[25] The conditions under which the investment function is monotonically increasing in productivity, which makes it possible to invert the investment function and gives an expression for productivity as a function of capital and investment, can be found in Olley and Pakes (1996).

time-invariant industry-specific production function coefficients. So within-industry variation is used to identify the coefficients of equation (11). Although this first regression provides consistent estimates for β_{0i}, β_{bl}, β_{wh}, and β_M, it does not identify the capital coefficients β_K and β_S.

The second regression, equation (12), focuses on the probability of a firm's survival, estimated using independent Logit functions for the pre-1991 and the post-1991 data, taking into account that the shutdown probabilities may have changed systematically after trade liberalization. Muendler (2004a) estimates probabilities over a fourth-order polynomial in $(I_{it}^k, I_{it}^s, a_{it}, k_{it}, s_{it}, k_{it}^F)$ and $\mathbf{D}_t$.

$$\Pr(\chi_{it+1} = 1|.) = P\left(I_{it}^k, I_{it}^s, a_{it}, k_{it}, s_{it}, k_{it}^F; D_{it}\right) \quad (12)$$

A third-order polynomial expansion approximates the expectation of a survivor's productivity ω_{it+1} one period in advance, as shown below, in equation (13)

$$\sum_{m=0}^{3} \sum_{n=0}^{3-m} \beta_{mn} (\hat{P})^m (\hat{h})^n \approx \int_{\underline{\omega}(k_{it}, s_{it}, \boldsymbol{D}_t)} \omega_{it+1} \frac{f(\omega_{it+1}|\omega_{it})}{\Pr(\chi_{it+1}=1|.)} d\omega_{it+1} \quad (13)$$

where $\underline{\omega}(k_{it}, s_{it}, \boldsymbol{D}_t)$ is the smallest productivity level that a firm with capital k_{it} and s_{it} needs in order to stay in business under market conditions $\mathbf{D}_t$. The $\hat{P}$ term in the polynomial expansion is the Logit-predicted survival likelihood. The unknown productivity component $\hat{h}$ results from $\hat{h}(.) = \hat{\phi}(.) - +\hat{\beta}_K k_{it+1} + \hat{\beta}_S s_{it+1}$. These considerations give rise to the third regression, equation (14). The equipment and structures coefficients (β_K and β_S) are estimated by non-linear least squares, using the estimates from firm fixed-effects regressions as starting values.

$$y_{it+1} - \hat{\beta}_{0i} - \hat{\beta}_{bl} l_{it+1}^{bl} - \hat{\beta}_{wh} l_{it+1}^{wh} - \hat{\beta}_M m_{it+1}$$
$$= \beta_K k_{it+1} + \beta_s s_{it+1} + \sum_{m=0}^{3} \sum_{n=0}^{3-m} \beta_{mn} (\hat{P})^m (\hat{h})^n + \eta_{it+1} \quad (14)$$

Finally, with all the estimated coefficients, the ln OP-TFP at the firm level is given by

$$\widehat{lnOP-TFP}_{it} = \mathrm{y_{it}} - \bar{\beta}_J - \hat{\beta}_{bl} l_{it}^{bl} - \hat{\beta}_{wh} l_{it}^{wh} - \hat{\beta}_M m_{it} - \hat{\beta}_K k_{it} - \hat{\beta}_s s_{it}$$

where $\bar{\beta}_J$, the average firm fixed effect, is defined as $\bar{\beta}_J \equiv \sum_{j \in S}^{J} \beta_{oj}/J$ (J is the number of firms in the industry), which eliminates confounding time-invariant demand conditions from ln *TFP-EOP*.

Table 1 – Correlation across TFP measures

Correlation\Sample	OLS-TFP All	OP-TFP All	OLS-TFP Stayers	OP-TFP Stayers
OLS-TFP All	1			
OP-TFP All	0.921	1		
OLS-TFP Stayers	0.999	0.9172	1	
OP-TFP Stayers	0.9214	0.9989	0.9193	1

All TFP variables are expressed in natural logarithms. OLS-TFP is the TFP obtained through the use of an OLS estimator, while OP-TFP is obtained by the Muendler (2004a,b) extended Olley-Pakes methodology. The "All" sample is an unbalanced panel of firms. The "Stayers" sample is a balanced panel of firms that were active throughout the trade liberalization period.

Table 2 – Descriptive Statistics

Variable	Number of Observations	Mean	Std. Dev.	Minimum	Maximum
OLS-TFP All	160	0.9849	0.0375	0.9127	1.0973
OP-TFP All	160	0.9981	0.0307	0.9177	1.0781
OLS-TFP Stayers	160	0.9851	0.0376	0.9134	1.0975
OP-TFP Stayers	160	0.9983	0.0307	0.9181	1.0781
Brazilian output tariff	160	0.2031	0.1212	0.040	0.750
Brazilian input tariff	160	0.2193	0.0998	0.0914	0.6818
Colombian output tariff	160	0.3193	0.1850	0.0649	1.199

All TFP variables are expressed in natural logarithms. OLS-TFP is the TFP obtained through the use of an OLS estimator, while OP-TFP is obtained by the Muendler (2004a, b) extended Olley-Pakes methodology. The "All" sample is an unbalanced panel of firms. The "Stayers" sample is a balanced panel of firms that were active throughout the trade liberalization period. Input tariffs are calculated according to equation (1).

Table 3 – Baseline specification, equation (2), estimated by ordinary least squares (OLS)

	Dependent variable					
	OLS-TFP	OLS-TFP	OLS-TFP	OP-TFP	OP-TFP	OP-TFP
Independent variable	(1)	(2)	(3)	(4)	(5)	(6)
	0.054*	-0.113**	-0.109**	0.001	-0.110**	-0.107**
Output Tariff	(0.032)	(0.050)	(0.050)	(0.026)	(0.045)	(0.044)
	0.002	0.003		0.001	0.003	
Upstream TFP	(0.003)	(0.002)		(0.003)	(0.002)	
			0.011			0.022
Input Tariff			(0.057)			(0.048)
				0.001	0.003	
Year and industry dummies	No	Yes	Yes	No	Yes	Yes
R^2	0.0329	0.780	0.778	0.00107	0.777	0.775
Observations	160	160	160	160	160	160

All TFP variables are expressed in natural logarithms. Robust standard errors reported in parentheses.
**, * indicates statistical significance at the 5% and 10% levels, respectively.

Table 4 – Baseline specification, equation (2), estimated by instrumental variables

	Dependent variable					
	OLS-TFP	OLS-TFP	OLS-TFP	OP-TFP	OP-TFP	OP-TFP
Independent variables	(1)	(2)	(3)	(4)	(5)	(6)
Output Tariff	-0.239**	-0.383**	-0.383**	-0.187**	-0.328**	-0.328**
	(0.078)	(0.102)	(0.101)	(0.063)	(0.085)	(0.085)
Upstream TFP	0.004*		0.005	0.003		0.004
	(0.002)		(0.003)	(0.002)		(0.002)
Input Tariff		0.326**	0.324**		0.318**	0.316**
		(0.104)	(0.103)		(0.087)	(0.087)
Observations	160	160	160	160	160	160
1st Stage						
Colombian Tariff	0.281**			0.281**		
	(0.053)			(0.053)		
Upstream TFP	0.007			0.007		
	(0.006)			(0.005)		
Weak id. Kleibergen-Paap rk Wald F statistic	28.58	12.10	12.14	28.58	12.10	12.15
Stock-Yogo 10% max IV relative bias critical values	16.38	7.03	7.03	16.38	7.03	7.03
Endogeneity test	3.429*	28.97**	29.04**	1.929	35.12**	35.31**
	[0.064]	[0.000]	[0.000]	[0.165]	[0.000]	[0.000]

All TFP variables are expressed in natural logarithms. Year and industry effects included in all specifications. Robust standard errors are reported in parentheses. *p*-values are reported in brackets. **, * indicates statistical significance at the 5% and 10% levels, respectively. Endogeneity test null hypothesis is that import tariff and input import tariff (if included in the estimated specification) are exogenous regressors.

Table 5 –First difference of the baseline specification, equation (2), estimated by instrumental variables

	Dependent variable					
	OLS-TFP	OLS-TFP	OLS-TFP	OP-TFP	OP-TFP	OP-TFP
Independent variables	(1)	(2)	(3)	(4)	(5)	(6)
ΔOutput Tariff	-0.197**	-0.194**	-0.189**	-0.141**	-0.144*	-0.139*
	(0.074)	(0.095)	(0.093)	(0.058)	(0.075)	(0.073)
ΔUpstream TFP	0.002		0.002	0.001		0.001
	(0.001)		(0.001)	(0.001)		(0.001)
ΔInput Tariff		-0.020	-0.022		-0.003	-0.005
		(0.116)	(0.114)		(0.091)	(0.090)
Observations	160	160	160	160	160	160
1st Stage						
ΔColombian Tariff	0.331**			0.332**		
	(0.056)			(0.056)		
ΔUpstream TFP	0.004*			0.004		
	(0.003)			(0.003)		
Weak id. Kleibergen-Paap rk Wald F statistic	34.84	8.985	9.387	34.84	8.985	9.390
Stock-Yogo 10% max IV relative bias critical values	16.38	7.03	7.03	16.38	7.03	7.03
Endogeneity test	2.728*	3.227	3.013	1.647	1.371	1.182
	[0.099]	[0.199]	[0.222]	[0.199]	[0.504]	[0.554]

All TFP variables are expressed in natural logarithms. Year dummy variables included in all specifications. Colombian tariffs and input tariffs calculated using Colombian tariffs (if Brazilian input tariffs included in the estimated specification), which are used as excluded instruments. Robust standard errors are reported in parentheses. *p*-values are reported in brackets. **, * indicates statistical significance at the 5% and 10% levels, respectively. Endogeneity test null hypothesis is that import tariff and input import tariff (if included in the estimated specification) are exogenous regressors.

Table 6 – Estimates of the augmented baseline specification, equation (3), using the Arellano-Bond estimator to account for TFP time persistence.

	Dependent variable					
	OLS-TFP	OLS-TFP	OLS-TFP	OP-TFP	OP-TFP	OP-TFP
Independent variables	(1)	(2)	(3)	(4)	(5)	(6)
ΔTime Lag TFP	0.202**	-0.041	-0.040	0.370**	0.150	0.162
	(0.099)	(0.154)	(0.138)	(0.104)	(0.175)	(0.149)
ΔOutput Tariff	-0.016	-0.214**	-0.113*	-0.049	-0.282**	-0.153**
	(0.043)	(0.091)	(0.067)	(0.036)	(0.108)	(0.070)
ΔUpstream TFP	0.001	0.002	0.002	0.001	0.002	0.002
	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)
ΔInput Tariff			0.095			0.014
			(0.141)			(0.126)
Overidentification test	59.19**	30.30**	44.69**	47.47**	15.73	29.05**
	[0.000]	[0.017]	[0.000]	[0.006]	[0.472]	[0.024]
Test for AR(1) error	-3.945**	-0.528	-0.697	-4.026**	-0.857	-1.480
	[0.000]	[0.598]	[0.486]	[0.000]	[0.391]	[0.139]
Test for AR(2) error	0.007	-1.036	-0.810	-0.155	-1.465	-1.045
	[0.995]	[0.300]	[0.418]	[0.877]	[0.143]	[0.296]
Arellano-Bond instruments lags used	2 to 4	3 to 4	3 to 4	2 to 4	3 to 4	3 to 4
Observations	160	160	160	160	160	160

All TFP variables are expressed in natural logarithms. Colombian tariffs and input tariffs calculated using Colombian tariffs (if Brazilian input tariffs included in the estimated specification), which are used in addition to the traditional Arellano-Bond instruments. Year dummies included in all regressions. Standard errors are reported in parentheses. *p*-values are reported in brackets. **, * indicates statistical significance at the 5% and 10% levels, respectively.

Table 7 – TFP spillovers accounted for by estimating equation (5) using the Generalized Spatial Two Stage Least Square estimator.

	Dependent variable					
	OLS-TFP	OLS-TFP	OLS-TFP	OP-TFP	OP-TFP	OP-TFP
Independent variables	(1)	(2)	(3)	(4)	(5)	(6)
Spatial lag ΔTFP	0.691**	0.551**	0.554*	0.783**	0.706**	0.667**
	(0.318)	(0.268)	(0.287)	(0.275)	(0.238)	(0.245)
ΔOutput Tariff	-0.173**	-0.152**	-0.152**	-0.122**	-0.108**	-0.093*
	(0.058)	(0.052)	(0.063)	(0.045)	(0.042)	(0.049)
ΔInput Tariff			0.002			-0.072
			(0.160)			(0.120)
Error Spatial correlation		-0.451	-0.456		-0.480	-0.465
		(0.310)	(0.313)		(0.301)	(0.318)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	160	160	160	160	160	160

All TFP variables are expressed in natural logarithms. Colombian tariffs and input tariffs calculated using Colombian tariffs (if Brazilian input tariffs included in the estimated specification), which are used in addition to Kelejian and Prucha (1998) instruments. Standard errors are reported in parentheses. **, * indicates statistical significance at the 5% and 10% levels, respectively.

Table 8 – Spillovers and time persistence of TFP accounted for by estimating equation (6) using the spatial Arellano-Bond estimator.

	Dependent variable					
	OLS-TFP	OLS-TFP	OLS-TFP	OP-TFP	OP-TFP	OP-TFP
Independent variables	(1)	(2)	(3)	(4)	(5)	(6)
Time Lag ΔTFP	0.210**	-0.006	-0.020	0.372**	0.190	0.153
	(0.105)	(0.148)	(0.161)	(0.109)	(0.159)	(0.183)
Spatial Lag ΔTFP	0.764**	0.725**	0.791***	0.564**	0.694**	0.798**
	(0.220)	(0.221)	(0.258)	(0.238)	(0.272)	(0.325)
ΔOutput Tariff	-0.051	-0.136*	-0.129	-0.065*	-0.178**	-0.160*
	(0.047)	(0.073)	(0.080)	(0.039)	(0.075)	(0.087)
ΔInput Tariff			-0.119			-0.148
			(0.180)			(0.166)
Overidentification test	40.46**	26.12*	22.15	38.72*	17.16	12.98
	[0.035]	[0.052]	[0.104]	[0.052]	[0.375]	[0.604]
Test for AR(1) error	-3.181**	-0.986	-1.104	-3.237**	-1.247	-1.351
	[0.001]	[0.324]	[0.270]	[0.001]	[0.212]	[0.177]
Test for AR(2) error	0.011	-0.375	0.0742	-0.244	-0.809	0.0652
	[0.991]	[0.707]	[0.941]	[0.807]	[0.418]	[0.948]
Arellano-Bond instruments lags used	2 to 4	3 to 4	3 to 4	2 to 4	3 to 4	3 to 4
Observations	160	160	160	160	160	160

All TFP variables are expressed in natural logarithms. Colombian tariffs and input tariffs calculated using Colombian tariffs (if Brazilian input tariffs included in the estimated specification), which are used in addition to the traditional Arellano-Bond instruments. Year dummies included in all regressions. Standard errors are reported in parentheses. *p*-values are reported in brackets. **, * indicates statistical significance at the 5% and 10% levels, respectively.

Table 9 – TFP spillovers accounted for by estimating equation (5) using the Generalized Spatial Two Stage Least Square estimator with heteroskedastic robust standard errors.

	Dependent variable					
	OLS-TFP	OLS-TFP	OLS-TFP	OP-TFP	OP-TFP	OP-TFP
Independent variables	(1)	(2)	(3)	(4)	(5)	(6)
Spatial lag ΔTFP	0.691**	0.578**	0.588**	0.783**	0.724**	0.687**
	(0.284)	(0.238)	(0.253)	(0.256)	(0.214)	(0.206)
ΔOutput Tariff	-0.173***	-0.154**	-0.157**	-0.122**	-0.110**	-0.095*
	(0.061)	(0.053)	(0.066)	(0.051)	(0.045)	(0.051)
ΔInput Tariff			0.012			-0.072
			(0.140)			(0.087)
Error Spatial correlation		-0.497**	-0.498**		-0.486**	-0.467**
		(0.220)	(0.219)		(0.196)	(0.197)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	160	160	160	160	160	160

All TFP variables are expressed in natural logarithms. Colombian tariffs and input tariffs calculated using Colombian tariffs (if Brazilian input tariffs included in the estimated specification), which are used in addition to Kelejian and Prucha (1998) instruments. Robust standard errors calculated according to Kelejian and Prucha (2007) are reported in parentheses. **, * indicates statistical significance at the 5% and 10% levels, respectively.

Table 10 –First difference of the baseline specification, equation (2), estimated by instrumental variables (IV) with controls for market competitiveness.

	Dependent variable					
	OLS-TFP	OLS-TFP	OLS-TFP	OP-TFP	OP-TFP	OP-TFP
Technique	OLS	IV	IV	OLS	IV	IV
Independent variables	(1)	(2)	(3)	(4)	(5)	(6)
ΔOutput Tariff	-0.084*	-0.188**	-0.189*	-0.072**	-0.137**	-0.140*
	(0.043)	(0.090)	(0.104)	(0.033)	(0.069)	(0.081)
Herfindahl * ΔOutput Tariff	-0.000	-0.115	-0.116	0.009	-0.068	-0.068
	(0.054)	(0.195)	(0.196)	(0.035)	(0.150)	(0.151)
ΔInput Tariff	-0.031		0.001	-0.047		0.010
	(0.051)		(0.118)	(0.037)		(0.091)
Observations	160	160	160	160	160	160
Weak id. Kleibergen-Paap rk Wald F statistic		4.127	2.565		4.127	2.565
Stock-Yogo 10% max IV relative bias critical values		7.03	n.a.		7.03	n..a.
Endogeneity test		5.046*	4.927		3.028	2.443
		[0.080]	[0.177]		[0.220]	[0.486]

All TFP variables are expressed in natural logarithms. Year dummy variables included in all specifications. Robust standard errors are reported in parentheses. *p*-values are reported in brackets. **, * indicates statistical significance at the 5% and 10% levels, respectively. "n.a." indicates that the Stock-Yogo reference values do not exist.

Table 11 – TFP spillovers accounted for by estimating equation (5) using the Generalized Spatial Two Stage Least Square estimator with heteroskedastic robust standard errors and controls for market competitiveness.

	Dependent variable			
	OLS-TFP	OLS-TFP	OLS-TFP	OP-TFP
Independent variables	(1)	(2)	(3)	(4)
Spatial lag ΔTFP	0.676**	0.546**	0.711**	0.683**
	(0.295)	(0.248)	(0.251)	(0.214)
ΔOutput Tariff	-0.089	-0.103	-0.025	-0.039
	(0.112)	(0.092)	(0.094)	(0.079)
Herfindahl * ΔOutput Tariff	-0.135	-0.080	-0.121	-0.088
	(0.111)	(0.100)	(0.088)	(0.079)
ΔInput Tariff	-0.137	-0.116	-0.219	-0.189
	(0.215)	(0.182)	(0.145)	(0.130)
Error Spatial correlation		-0.454*		-0.357*
		(0.232)		(0.197)
Year dummies	Yes	Yes	Yes	Yes
Observations	160	160	160	160

All TFP variables are expressed in natural logarithms. Colombian tariffs and input tariffs calculated using Colombian tariffs (if Brazilian input tariffs included in the estimated specification), which are used in addition to Kelejian and Prucha (1998) instruments. Robust standard errors calculated according to Kelejian and Prucha (2007) are reported in parentheses. **, * indicates statistical significance at the 5% and 10% levels, respectively.