Openness to trade and industry cost dispersion: Evidence from a panel of Italian firms*

Massimo Del Gatto
University of Cagliari
and
CRENoS

Gianmarco I.P. Ottaviano
University of Bologna
FEEM
and
CEPR

Marcello Pagnini
Bank of Italy

May 4, 2007

Abstract

We use Italian firm-level data to investigate the impact of trade openness on the distribution of firms across marginal cost levels. In so doing, we implement a procedure that allows us to control not only for the standard transmission bias identified in firm-level TFP regressions but also for the omitted price bias due to imperfect competition. We find that more open industries are characterized by a smaller dispersion of costs across active firms. Moreover, in those industries the average cost is also smaller.

Keywords: Cost dispersion, openness to trade, firm-level data, firm selection, total factor productivity
J.E.L. Classification: F12, F15, R13.

*Corresponding author: Gianmarco I.P. Ottaviano, Università di Bologna, Dipartimento di Scienze Economiche, Piazza Scaravilli 2, 40126 Bologna, Italy. Email: gianmarco.ottaviano@unibo.it. We thank participants to seminars at the Bank of Italy and the University of Cagliari as well as participants to the conference "Empirical methods for the study of economic agglomeration" at Kyoto University for helpful comments. We are especially indebted with Eckhardt Bode, Pierre-Philippe Combes, Gilles Duranton, Thierry Mayer, Tomoya Mori, Henry Overman, David Weinstein and anonymous referees for useful comments. Ottaviano gratefully acknowledges financial support from MIUR and the European Commission.

1 Introduction

All industries are characterized by the simultaneous presence of firms with different levels of efficiency. More productive firms are larger in terms of output as well as sales and this maps into larger profits. Less productive firms are smaller and less profitable. More productive firms are also able to sell to more distant markets whereas less productive ones are confined to their local markets.

In this respect, a hallmark result goes under the label of "exceptional export performance" (Bernard and Jensen, 1999), which refers to the fact that exporters are systematically found to be on average more productive than purely local firms. In principle causality could run both ways: only more productive firms become exporters ('selection into export status') and exporting improves firm efficiency ('learning by exporting'). The current consensus view favors the former direction of causality. In particular, two stylized facts are often stressed. First, exposure to trade forces the least productive firms to exit the market or to shut down (Clerides et al., 1998; Bernard and Jensen, 1999; Aw et al., 2003). Second, trade liberalization leads to market share reallocations towards more productive firms (Pavcnik, 2002; Bernard et al, 2006). Thus, there seems to be some robust evidence that the opening of distant markets gives an additional opportunity only to more productive firms. This allows them to enlarge their markets shares to the detriment of less productive competitors, the least efficient of which are forced to exit.

These empirical results have been recently explained by theoretical models in which heterogeneous firms act in imperfectly competitive markets, perfect competition being incompatible with the equilibrium coexistence of firms with different efficiency levels. Existing models differ in terms of the feature that leads only the most productive firms to engage in distant trade. For example, Bernard et al. (2003) stress the role of limited product differentiation resulting in tougher worldwide price competition when markets become more open. Melitz (2003) builds instead on the role played by the sunk cost of exporting documented by Roberts and Tybout (1997) as well as Bernard and Jensen (1999). In the presence of these costs only more productive firms can afford the commercial opportunities of distant trade. When markups are fixed, this allows them to grow in size bidding up factor prices, which crowds out their less productive competitors. Melitz and Ottaviano (2005) show that a similar effect can derive from falling markups when these are allowed to react to increasing openness. Their model is extended and calibrated by Del Gatto et al. (2006) in the case of a multi-location multi-sector economy where the intensity of firm selection due to distant trade is shown to vary across locations depending on their sectoral specialization and their geographical position in the trade network.

All these models yield empirical predictions concerning the effects of trade openness on the distribution of firms across productivity or cost levels.[1] In particular, much attention has been devoted to the mean value of such distribution. For example, all models predict an increase in average industry productivity or a reduction in average industry cost as the market becomes more open to distant sellers and less productive

[1] This literature typically studies the effects of trade openness on Hicks-neutral technological change. While this remains also our focus, the potential impact of trade openness on relative factor intensity is an interesting topic for future research.

firms are forced to shut down. In Del Gatto et al. (2006) the average industry cost is lower in locations that have larger local markets, better access to distant markets and a bias towards industries with weaker product differentiation. On the other hand, less attention has been devoted to the dispersion (or spread) of the distribution. For example, the exit of the least productive firms should naturally lead to a reduction in the industry productivity or cost ranges, defined as the gaps between the best and the worst performing firms (or plants) in the industry. Accordingly, in a time-series perspective, increasing trade openness should be associated with falling productivity or cost ranges. In a cross-section perspective, smaller ranges should characterize larger and more accessible markets or sectors that are more open to distant trade and deal in less differentiated products. None of these implications for productivity or cost spreads have been explicitly tested so far. The only notable exception is Syverson (2004) who investigates the relation between productivity spreads and tradability as a robustness check within his analysis of the links between those spreads and product substitutability. His regressions, based on a cross-section of US firms, do not provide any support to the statistical significance of such relation.

The aim of this paper is to start filling the gap in the literature by investigating the theoretical implications of trade openness on industry cost dispersion. Specifically, we check whether *intra-industry cost spreads are smaller in more export-oriented industries.* This may have important implications for the political economy of trade liberalization. When firm heterogeneity is not taken into account, traditional trade models show that all firms equally lose from trade liberalization to the advantage of consumers. All firms have therefore the same incentive to participate to protectionist lobbying. Recent models with firm heterogeneity show instead that, within the very same sector, the incentive to lobby varies across firms with different market performance, which translates into political economy outcomes that depend on the dispersion of such performance (Bombardini, 2005). Accordingly, firm heterogeneity affects the clout of protectionist stances and their translation into effective pressure on policy makers.[2] Whether trade openness increases or decreases the differences between firms then becomes crucial for the political sustainability of ongoing trade liberalization.

Our empirical analysis is implemented on a panel of Italian firms drawn from the Company Accounts Data Service ('Centrale dei Bilanci'), which reports the balance sheets of around 30.000 firms from 1983 to 1999. From a methodological point of view, we show that, when bringing the existing theoretical models with firm heterogeneity to data, it is crucial to consistently account for the implications of their imperfectly competitive market structure, especially when it comes to measuring firm-level productivity. In so doing, we depart from the standard approach that estimates the total factor productivity (TFP) of a firm as the Solow residual of a firm-level regression under the assumption of perfectly competitive product markets. This assumption, which contradicts all theoretical models, has little bearing when the focus of the empirical analysis is on average industry productivity as usual. It is not innocuous when the focus is, instead, on the

[2] Bombardini (2005) shows that accounting for firm size dispersion and associated differences in lobby participation shares explains a non-negligible fraction of the variation of protection across US sectors. Our theoretical framework does not model this political economy channel. Nevertheless, we will have to take it into account when we discuss the estimated correlation between trade openness and dispersion of firm performance in terms of underlying causal relations.

productivity spread. To solve this problem, we implement a new estimation procedure that takes imperfect competition into account in the wake of Melitz (2000). This way we find that industries that are more export-oriented or exhibit lower transport costs generally exhibit smaller cost dispersion. To the best of our knowledge, we are the first to provide a systematic analysis of the relation between the degree of trade openness and the second moment of the distribution of firm efficiency for a European country.

The paper is organized in six sections after the introduction. Section 2 presents the theoretical framework. Section 3 uses the framework to illustrate the implications of imperfect competition for productivity estimation and implements a procedure that takes those implications into account. Section 4 describes the dataset. Section 5 presents a benchmark specification closely derived from the theoretical framework and reports the corresponding results. Section 6 discusses several robustness checks. Section 7 concludes.

2 Theoretical model

The theoretical framework is based on the model by Melitz (2003) and its extension by Falvey, Greenaway and Yu (2004). Consider an industry operating in two locations, home and foreign, that are identical in all exogenous attributes except market sizes. For parsimony, in the description of the model we focus on the home location with the understanding that analoguous conditions hold for the foreign one.

The industry is monopolistic competitive and each producer supplies a variety of a horizontally differentiated good using capital and labor as inputs. There is an infinite number of ex-ante identical potential entrants. In both countries each entrant faces the same sunk entry cost $r^{\alpha}w^{\beta}f_E$, where r and w are the (exogenously given) rental rate of capital and the wage of labor with complementary shares $\alpha + \beta = 1$. This cost can be thought of as an irreversible R&D investment whereby the entrant creates a new product and the corresponding production process. The outcome of R&D is uncertain in that the entrant gets to know the costs of its new production process only after entry has taken place and the corresponding cost has been sunk. All potential entrants make their decisions to enter or not simultaneously. Accordingly, after sinking their entry costs, all entrants are able to calculate the profits they would earn if they decided to stay in the market and start producing. This would require them to pay additional fixed costs of production $r^{\alpha}w^{\beta}f_D$ for domestic sales and $r^{\alpha}w^{\beta}f_X$ for distant sales ('exports') with $f_X > f_D$ as distant activities typically involve more complex operations. Distant sales also involve per-unit trade costs of the iceberg type: for one unit to reach distant customers $\tau > 1$ units have to be shipped.

All the foregoing costs are the same for all entrants in both countries and known before entry. The marginal cost is, instead, unknown before entry and thus may end up being different across entrants. Specifically, let C_i be a random variable with c.d.f. $G(C_i)$ over the support $[0, C_M]$. Then, the marginal cost is also a random variable and is assumed to be given by:

$$MC_i = r^{\alpha}w^{\beta}C_i \tag{1}$$

where C_i is the marginal resource cost (henceforth, 'real marginal cost' or RMC), which drives firm heterogeneity and is the inverse of TFP. As entrants become eventually different only after drawing their RMC's, it is convenient to identify them by the corresponding

RMC draws. Hence, we call i the entrant that draws a RMC realization equal to C_i. Accordingly, (1) is the marginal cost of entrant i.

In each location firms face the same downward sloping demands with constant own- and cross-price elasticities. In particular, using an asterisk to label foreign variables, entrant i in the home country faces demands

$$D_i = \frac{P_i^{-\sigma}}{P^{1-\sigma}} E \text{ and } \widetilde{D}_i = \frac{\widetilde{P}_i^{-\sigma}}{P^{*1-\sigma}} E^* \tag{2}$$

in its domestic and distant markets respectively. In (2), $\sigma > 1$ is the constant own- and cross-price elasticity of demand, P_i and $\widetilde{P}_i$ are the delivered prices in the domestic and foreign markets, E is industry expenditures as well as revenues, and P is the industry exact CES price index defined as

$$P = \left[\int_{i \in I} P_i^{1-\sigma} di + \int_{i \in I_X} P_i^{*1-\sigma} di \right]^{\frac{1}{1-\sigma}}$$

where I is the set of domestic producers while I_X is the set of foreign exporters.

Profit maximization yields equilibrium prices as constant markups over marginal costs:

$$P_i = \frac{\sigma}{\sigma - 1} r^\alpha w^\beta C_i \text{ and } \widetilde{P}_i = \tau P_i \tag{3}$$

which implies profits

$$\pi(C_i) = BC_i^{1-\sigma} \left(r^\alpha w^\beta\right)^{1-\sigma} - f_D r^\alpha w^\beta \text{ and } \widetilde{\pi}(C_i) = B^* \tau^{1-\sigma} C_i^{1-\sigma} \left(r^\alpha w^\beta\right)^{1-\sigma} - f_X r^\alpha w^\beta \tag{4}$$

in the domestic and distant markets respectively with $B \equiv EP^{\sigma-1}[(\sigma - 1)/\sigma]^{\sigma-1}/\sigma$. Then, only entrants with RMC weakly below the cutoffs C_D and C_X, such that $\pi(C_D) = 0$ and $\pi(C_X) = 0$, are able to serve respectively only the local market and also the distant market without making losses. Other entrants do not produce and exit. Given (4), $\pi(C_D) = 0$ and $\pi(C_X) = 0$ can be expressed as

$$B = f_D C_D^{\sigma-1} \left(r^\alpha w^\beta\right)^\sigma \text{ and } B^* = f_X \tau^{\sigma-1} C_X^{\sigma-1} \left(r^\alpha w^\beta\right)^\sigma \tag{5}$$

hence the cutoffs for local sales and for exports to the same market are linked by the following relation

$$C_X^* = \left(\frac{f_X}{f_D} \frac{1}{\tau^{1-\sigma}} \right)^{\frac{1}{1-\sigma}} C_D \tag{6}$$

This shows that larger fixed and per-unit trade costs increase the gap between the domestic and export RMC thresholds.

All this being anticipated by prospective entrants, free entry implies that the expected profits from production exactly match the entry cost:

$$\int_0^{C_D} \pi(C_i) dG(C_i) + \int_0^{C_X} \widetilde{\pi}(C_i) dG(C_i) = f_E r^\alpha w^\beta \tag{7}$$

After substituting (4) and (5), expressions (6) and (7) together with their analogues for the foreign location define a system of four equations in the four unknown cutoffs.

All results so far have been derived for a generic distribution of RMC draws. To bring the theoretical model to data, however, we need to make some specific assumption on the functional form of such distribution. In particular, we assume that entrants draw their RMC's from a Pareto distribution. In so doing, we build on Del Gatto et al. (2006) who find that a Pareto c.d.f. provides a fairly accurate description of actual data across many industries in EU countries. While this is our working assumption, in the empirical analysis we will test its validity and also investigate a more general specification where such assumption is removed.[3] Formally, let us assume that entrants draw their RMC's from the following c.d.f.:

$$G(C_i) = \left(\frac{C_i}{C_M}\right)^k \text{ for } C_i \in [0, C_M] \tag{8}$$

where $k \geq 1$ measures the skewness of the RMC distribution towards the upper bound of its support. Then, under the regularity condition $k > \sigma - 1$, the domestic cutoff is the same in both locations and evaluates to

$$C_D = \left[\frac{f_E(\Lambda - 1)}{f_D \Omega}\right]^{\frac{1}{k}} C_M \tag{9}$$

where $\Lambda = k/(\sigma - 1) > 1$ and $\Omega = 1 + \tau^{-k} (f_X/f_D)^{-(\Lambda-1)}$ is a synthetic measure of trade openness (Baldwin, 2005). This index summarizes the influence of variable and fixed trade costs as mediated by the degree of product substitutability σ and the shape parameter k of the Pareto distribution. According to (9), freer trade (smaller τ or smaller f_X) is associated with a smaller domestic cutoff C_D and, by (6), with a larger export cutoff C_X. Under the Pareto assumption, that maps into a lower (unweighted) mean RMC of local sellers:

$$\text{Mean} = \frac{k}{k+1} C_D \tag{10}$$

The mean, however, is not a good measure of central rendency in skewed distributions like the Pareto for $k > 1$. For such distributions a better measure is the median, which under our Pareto assumption is:

$$\text{Median} = \left(\frac{C_D}{2}\right)^{\frac{1}{k}} \tag{11}$$

Both (10) and (11) reveal the essence of the 'selection effect' highlighted by Melitz (2003): more openness to trade translates into tougher competition in factor markets, which leads to a lower domestic cutoff. Firms with RMC above this level are forced to leave the market and their shares are reallocated towards more efficient firms. Hence we have:

Result 1. *Lower trade costs are associated with a lower central tendency of marginal costs.*

[3] The assumption that RMC's are drawn from a Pareto distribution is very convenient in theoretical models when it comes to deriving closed form solutions. Which properties of those solutions also hold for general classes of distributions is an important direction of future research. See, e.g., Combes at al (2007).

Turning to cost dispersion, its simplest measure is the 'range', defined as the RMC gap between the best and the worst performing firms. Given the support $[0, C_D]$, the range is simply equal to C_D, so it is smaller when trade is freer. While easily understood, being based on two values only, the range is very sensitive to extreme observations and should be used together with other measures. The most widely used measure of dispersion is the 'standard deviation', which for our Pareto distribution evaluates to:

$$\text{Std.Dev.} = \sqrt{\frac{k}{(k+1)^2 (k+2)}} C_D \tag{12}$$

which is an increasing function of the cutoff C_D and thus also smaller when trade is freer. The standard deviation, however, is not a good measure of dispersion in skewed distributions, for which a better measure is the 'interquartile range' defined as the difference between the 75th and the 25th percentiles. Under our Pareto assumption, the interquartile range is:

$$\text{IQ Range} = \left[(0.75)^{\frac{1}{k}} - (0.25)^{\frac{1}{k}}\right] C_D \tag{13}$$

Also this measure is increasing in the cutoff C_D, thus showing that dispersion is lower when openness is larger. We can therefore conclude with:

Result 2. *Lower trade costs are associated with a smaller dispersion of marginal costs.*

Due to the general skewness of the Pareto distribution, in the empirical analysis we will choose the IQ Range as our preferred dispersion measure and the median as our preferred central tendency measure while investigating alternative measures as robustness checks.[4]

3 Productivity estimation

To calculate the individual RMC's we rely on the production function associated with the marginal cost (1):

$$Y_i = A_i (K_i)^{\alpha} (L_i)^{\beta} \tag{14}$$

K_i and L_i are capital and labor inputs respectively while A_i is TFP such that $A_i = 1/C_i$. Hence, RMC's can be retrieved from data by inverting the corresponding estimated TFP's.

The most common approach to (individual) TFP estimation expresses (14) in logs (in lowercase letters) and includes a multiplicative stochastic disturbance ($e^{u_{it}}$):

$$y_{it} = a_{it} + x_{it}\theta + u_{it} \tag{15}$$

[4]Results 1 and 2 have been derived under the assumption that RMC's are drawn from a Pareto distribution. The underlying logic, however, applies to a larger class of distributions. In Sections 5 we will therefore use a measure of the IQ range that is not based on the Pareto assumption. Moreover, in Sections 6.2 and 6.3 we will test alternative dispersion measures as well as a more general empirical specification of the relation between trade openness and cost dispersion.

In equation (15) a_{it} is individual productivity, x_{it} is the vector of inputs, θ is the vector of the elasticities of output with respect to each input and u_{it} is meant to capture measurement errors as well as unobserved firm-specific shocks. The value of a_{it} can then be recovered by estimating the vector $\hat{\theta}$, computing the fitted value of firm i's output $\hat{y}_{it}$ and deriving $\hat{a}_{it}$ as the (exponential of the) difference between y_{it} and $\hat{y}_{it}$ ('Solow residual'). However, a word of caution on the estimation procedure is in order.[5]

3.1 Omitted price bias and transmission bias

Standard OLS estimation of (15) could run in two different problems. The first problem is often referred to as 'simultaneity' and stems from the fact that a_{it} is 'unobservable' to the econometrician but it is reasonably observed by the firm and likely to influence its input choice. This relegates a_{it} to the error term of (15), which is therefore correlated with the vector of inputs x_{it} biasing the OLS estimates. In econometric parlance, a_{it} is said to 'transmit' to the explanatory variables, hence the term 'transmission bias'. The second problem originates from the unavailability of information on firms' output y_{it} in our dataset. This forces us to use a proxy consisting of sales deflated by an industry-wide price-index, given that prices are themselves not available at firm level. Such circumstance has no relevance under perfect competition as all firms quote the same price. On the contrary, when markets are imperfectly competitive, firm-level estimated productivity is likely to be misstated. This, of course, applies to our monopolistic competition model, in which, as highlighted by equation (3), differences in productivity entirely translate into differences in prices as these are determined by markups over marginal costs. Since the problem is caused by omitting the individual price from the estimation, this problem is usually referred to as 'omitted price bias'.

The two problems can be summarized following Klette and Griliches (1996) in rewriting the estimating version of equation (15) as follows:

$$\tilde{r}_{it} = x_{it}\theta^r + u^r_{it} \tag{16}$$

where $u^r_{it} = a^r_{it} + e^r_{it}$ and, due to constrained data availability, physical output has been replaced by deflated sales $\tilde{r}_{it} = r_{it} - p_t = y_{it} + q_{it}$, with r_{it} indicating firm revenues and $q_{it} = p_{it} - p_t$ measuring the difference between (the log of) the firm specific price p_{it} and (the log of) the deflator p_t. Given M observations, in matrix notation the OLS estimator of θ^r is:

$$\hat{\theta}^r = \left(X'X\right)^{-1} X'\widetilde{r} \tag{17}$$

where $\widetilde{r}$ is the vector of deflated sales and X is the matrix of factor inputs. Since $\tilde{r}_{it} = y_{it} + q_{it}$, the probability limit of $\hat{\theta}^r$ is

$$plim_{N\to\infty}(\hat{\theta}^r) = \theta + plim_{N\to\infty}\left[\left(X'X\right)^{-1} X'q\right] + plim_{N\to\infty}\left[\left(X'X\right)^{-1} X'a\right] \tag{18}$$

where q is the vector of the differences between individual prices and the industry deflator, a is the vector of individual productivities in (15), and we assume orthogonality in the

[5]The estimation of production functions raises additional issues that go beyond the scope of the present paper. Among others, it is worthwhile mentioning functional form specification and input heretogeneity, which may be particularly relevant in the case of labor.

error term e^r_{it}. On the right hand side, the second and third terms are the omitted price and the transmission biases. They can be seen respectively as the OLS estimators of ω and ε in the auxiliary regressions $q = X\omega + u^q$ and $a = X\varepsilon + u^a$, where u^q and u^a are orthogonal error terms. Accordingly, we can write:

$$plim_{N\to\infty}(\hat{\theta}^r) = \theta + \omega + \varepsilon \tag{19}$$

so that in the limit estimated TFP evaluates to

$$plim_{N\to\infty}(\widehat{a}_{it}) = \widetilde{r}_{it} - x_{it}\, plim_{N\to\infty}(\hat{\theta}^r) = a_{it} - x_{it}\omega - x_{it}\varepsilon \tag{20}$$

where $x_{it}\omega$ and $x_{it}\varepsilon$ are the associated omitted price and transmission biases.

3.2 Dealing with the biases

We deal with the two types of bias as follows. We start with removing the transmission bias following the procedure suggested by Levinsohn and Petrin (2003) in the absence of the omitted price bias (henceforth LP). In loose terms, the transmission bias originates from the facts that the productivity shock a is observed by the firm but unobserved by the econometrician so that ε cannot be directly estimated. The solution consists in finding an observable proxy variable that 'reacts' to variations in a. Provided that the relationship between such proxy and the productivity shock can be inverted, its inversion can be used to 'internalize' the transmission bias and to identify the production coefficients consistently.[6] This approach, first proposed by Olley and Pakes (1996) using investment as a proxy, has been amended by Levinshon and Petrin (2003) who use instead intermediates and show that these can be introduced as additional inputs in the production function (15) to remove the transmission bias (i.e. $\varepsilon = 0$).[7] Under perfect competition, the LP estimator $\hat{\theta}^{LP}$ is consistent as $p_{it} = p_t$ implies $\omega = 0$ so that $plim_{N\to\infty}(\hat{\theta}) = \theta$ and $plim_{N\to\infty}\hat{a}^{LP}_{it} = a_{it}$.

Whenever $p_{it} \neq p_t$, the OLS estimator is, however, affected by the omitted price bias ($\omega \neq 0$). To solve this problem, we reinterpret the approach to the 'inconsistency of common scale estimators' suggested by Klette and Griliches (1996). In so doing, we follow Melitz (2000) who builds on that approach to derive the conditions under which the LP method still yields consistent estimates under monopolistic competition. We call LPM the resulting procedure.

Using (2) and (15), the ratio of the firm price to the industry deflator can be expressed as

$$q_{it} = p_{it} - p_t = \frac{1}{\sigma}\bar{\widetilde{r}}_t - \frac{1}{\sigma}(x_{it}\theta + a_{it}). \tag{21}$$

[6] The invertibility condition simply requires intermediate input use to be increasing in TFP conditional on capital. Melitz (2000) shows that, under monopolistic competition, such a monotonicity condition holds whenever more productive firms do not set disproportionately higher markups than less productive firms. Formally, this requires the elasticity of the markup with respect to productivity to be bounded above by $(\sigma - 1)/\sigma$.

[7] The LP procedure involves two stages. If invertible, the intermediates demand function $m_{it} = m(a_{it}, \cdot)$ is used to generate the proxy $a_{it} = g(m_{it}, \cdot)$. In the first stage, this is substituted in the production function in order to identify the labor parameter getting rid of the correlation between the regressors and the error term. In the second stage, the estimated value of the labor parameter is then used to recover the values of the other parameters.

where $\bar{\bar{r}}_t = [(r_t - p_t) - n_t]$ denotes average deflated sales. Equation (21) identifies the sources of the omitted price bias in our monopolistic competition framework and we can use it together with equation (15) to purge deflated sales of unobserved output:

$$\tilde{r}_{it} = \frac{1}{\sigma}\bar{\bar{r}}_t + \left(1 - \frac{1}{\sigma}\right) x_{it}\theta + \left(1 - \frac{1}{\sigma}\right)(a_{it} + e_{it}) \tag{22}$$

Consistent estimates of θ and σ can now be obtained by applying the LP procedure to (22). If we call them $\hat{\theta}^{LPM}$ and $\hat{\sigma}^{LPM}$ respectively, the LPM estimated productivity can be written as

$$\hat{a}_{it}^{LPM} = -x_{it}\hat{\theta}^{LPM} + \frac{1}{\hat{\sigma}^{LPM} - 1}\left(\bar{\bar{r}}_t - \hat{\sigma}^{LPM}\tilde{r}_{it}\right). \tag{23}$$

Since the difference q_{it} between the individual price and the industry deflator is now correctly identified, the estimated productivity in (23) measures the 'true' productivity (i.e. $\hat{a}_{it}^{LPM} = a_{it}$). We can, therefore, calculate the RMC of firm i as $\widehat{C}_i = \exp(-\hat{a}_{it}^{LPM})$.

Three comments are in order. First, with respect to the standard LP procedure assuming perfect competition, consistent estimation under monopolistic competition requires average deflated sales $\bar{\bar{r}}_t$ to appear as an additional regressor in the first stage. Second, the fact that regression (22) consistently estimates the elasticity of substitution σ is of interest in itself as it provides a new way of computing the extent of product differentiation across sectors. Third, to isolate the omitted price bias, note that neglecting price dispersion q_{it} in the LP regression would imply estimated productivity $\hat{a}_{it}^{LP} = \tilde{r}_{it} - x_{it}\hat{\theta}^{LP} = \hat{a}_{it}^{LPM} - x_{it}\omega$. Hence, we would have:

$$\hat{a}_{it}^{LP} = a_{it} - \frac{1}{2\left(\hat{\sigma}^{LPM} - 1\right)}(r_{it} - \bar{r}_t) \tag{24}$$

where $\bar{r}_t = r_t - n_t$ stands for average non-deflated sales and we have used the fact that $\hat{a}_{it}^{LPM} = a_{it}$ and $x_{it}(\hat{\theta}^{LPM} - \hat{\theta}^{LP}) = x_{it}\omega$ by (19). Hence, the 'correction factor' to be applied to the standard LP estimate $\hat{a}_{it}^{LP}$ is an increasing function of a firm's sales relative to the industry average, being positive for above average firms and negative for below average ones. In other words, *disregarding price dispersion results in understating the productivity of firms that are more productive (and therefore larger) than the average and overstating the productivity of firms that are less productive (and therefore smaller) than the average.*[8] Moreover, the magnitude of the bias depends on the estimated elasticity of substitution $\hat{\sigma}^{LPM}$: the lower the elasticity of substitution (i.e. the more differentiated the products), the larger the bias. Finally, the bias vanishes on average, so the estimated average productivity is unaffected by the correction. The same, however, does not apply to the dispersion measures.

4 Data description

Our data are drawn from 'Centrale dei Bilanci' (Company Accounts Data Service, henceforth CADS), which reports the balance sheets of around 30.000 Italian firms for the time

[8]This is consistent with the evidence recently reported by Foster et al. (2005), who have the rare chance of comparing for several industries the estimated productivity outcomes resulting from either firm output or firm deflated sales. Although they find that the two measures are highly correlated, they nonetheless show that quantity-based productivity measures exhibit greater dispersion than revenue-based ones.

span 1983-1999.[9] It is worth noting that, due to the specific focus of the archive, the sample is relatively skewed towards larger firms and business units located in northern Italy.[10] Balance sheet data encompass value added, the value of intermediate goods and services, the capital stock at book value and the number of employees. Further information refers to the location of firms down to the municipality level. The sectoral classification we adopt, which includes 18 sectors, is very similar to the two-digit Nace Rev.1 disaggregation and tries to achieve a good compromise between the need of homogeneity within a given sector and the need of a sufficient number of observations for a statistically reliable analysis. Not all the firms belonging to the CADS sample report the same balance sheet information. Specifically, only a subset of them reports the full set of balance sheet items that are used to estimate capital stocks as described in Section 5.2. Thus, we are forced to restrict the analysis to this group of firms with a consequent fall in the number of business units, especially small units, included in our sample. Tables 4 and 5 show the coverage in 1991 and 1996, obtained by comparing the CADS sample with the universe of Italian firms reported by the Italian NIS (i.e. ISTAT). As evident, our sample bias is much lower with respect to firms with more than 20 workers. This disadvantage is partially compensated by a better and more reliable quality of balance sheet information reported. We have excluded firms in the first and the last percentile in terms of factor shares and capital-to-labour ratio. We have also purged the sample of firms with intermittent participation and firms observed for one year only. The resulting numbers of firms are reported, respectively by sector and year of observation, in Tables 2 and 3, which also show the percentage of exporting firms (i.e. firms that served foreign markets in more than half of the period they were observed in the sample). Note that the percentage of exporters is rather high due to the large size of the firms included in the sample.

5 Benchmark specification

Our aim is to investigate the relation between RMC spreads and openness to trade as highlighted by Result 2 of the theoretical model in Section 2. Along the way, we will also check the empirical relevance of Result 1.

5.1 Estimating equation

Let us start by substituting the domestic cutoff cost (9) in the interquartile range dispersion measure (13). The logarithmic transformation of the resulting equation gives a downard sloping log-linear relation between RMC dispersion and trade openness Ω. Although this structural relation is industry specific, limited degrees of freedom call for some extra constraint on parameters. In particular, we assume that the 'slopes' are the same across industries and rely on fixed effects ('intercepts') to account for industry

[9] CADS was established in the early 1980s jointly by the Bank of Italy, the Italian Banking Association (ABI) and a pool of leading banks with the aim of collecting and sharing information on borrowers.

[10] For a discussion about the representativeness of the sample, see Cingano and Schivardi (2004).

differences. Our empirical specification then becomes:

$$\begin{aligned} disp_{st} &= \alpha + \beta omega_{st} + \sum_s \iota_s I_s + \sum_t \iota_t I_t + \varepsilon^d_{st} \\ &\quad \text{with} \quad s = 1, \ldots, 17 \quad t = 1983, \ldots, 1999 \end{aligned} \tag{25}$$

where, for sector s at time t, $disp_{st}$ is the (log) measure of RMC dispersion and $omega_{st}$ is the (log) index of trade openness Ω. A vector I_s of industry fixed effects, a vector I_t of time fixed effects, and the disturbance term ε^d_{st} complete the specification.

Three versions of (25) are presented. In Model [1] we run a simple bivariate regression of $disp_{st}$ on $omega_{st}$. This should not suffer from spurious time-series correlation as neither productivity dispersion nor trade openness exhibit monotone trends (see Tables 8 and 10). In Model [2] we introduce sectoral fixed effects to control for time-invariant industry characteristics that may impact on RMC dispersion. We also introduce year fixed effects to control for specific events that may affect the RMC distribution and are common to all sectors. In 1992, for instance, the Italian economy underwent a financial and economic crisis that had a strong impact on all manufacturing economic activities. In 1998-1999 some European countries, including Italy, moved to a fixed exchange regime with an impact on all industrial sectors.[11] Finally, in Models [1] and [2] the disturbance term ε^d_{st} is assumed to be iid. However, both $disp_{st}$ and $omega_{st}$ result from estimation procedures, hence potential measurement errors stemming from the corresponding regressions may lead to inefficient coefficient estimates. Industry and year fixed effects should reduce the impact of of measurement errors on our estimates, yet we cannot completely rule out that concern. For this reason, as suggested by Beck and Katz (1995), in Model [3] we re-estimate (25) using the Prais-Winsten method with panel corrected standard errors (PCSE) computed under the assumption of heteroskedastic disturbances across industries.[12]

5.2 Productivity measure

For the estimation of individual productivity, we have adjusted both value added and investment using the two-digit deflators provided by NIS's National Accounts. The capital stock at firm level has been obtained from the book value of investment using the permanent inventory method and accounting for the sector-specific depreciation rates from NIS's National Accounts data. The initial capital stock has been estimated using the deflated book value, adjusted for the average age of capital calculated from the depreciation fund. Descriptive statistics for key variables for the estimation of productivity are reported in Table 6.

TFP is estimated by the LPM procedure described in Section 3 using 'intermediate goods and services' as the LP proxy variable. The output shares of capital and labour are allowed to vary across sectors but not across years. Tables 7 and 8 respectively

[11] The implications of the theoretical framework have been derived under the assumption that factor prices are equalized and technologies are the same across locations. Industry and time fixed effects should also help to control for possible country asymmetries.

[12] Assuming PCSE is equivalent to assuming that disturbances are no longer iid. Alternatively, one could use feasible generalized least squares. Beck and Katz (1995) show, however, that PCSE performs much better when, as in our case, the time span is relatively small with respect to the size of the cross section.

report estimated average productivity by industry and by year together with various average dispersion measures (the interquartile range 'IQ Range', the interdecile range '90-10 Range', the Pareto skewness parameter 'k_s').[13] All dispersion measures reveal sharp differences across industries. For instance, TFP dispersion in sector 1 ('Food and beverages') is more than double the ones observed in sector 16 ('Furniture') and in sector 17 ('Jewelery and related articles, Musical instruments, Sport goods, Toys and Games'). Moreover, productivity dispersion shows no monotonic trend over time.

Tables 7 and 8 also report the decomposition of productivity growth following Pavcnik (2002). In a given year, the weighted average productivity of a sector ($\sum_i \eta_{it}\hat{a}_{it}^{LPM}$, where η_{it} is the share of firm i in sectoral output) is decomposed in two components: an unweighted average productivity measure ($\overline{\hat{a}}_t^{LPM}$) and a covariance term ($\sum_i(\eta_{it} - \overline{\eta}_t)(\hat{a}_{it}^{LPM} - \overline{\hat{a}}_t^{LPM}$), where $\overline{\eta}_t$ is average firm share in sectoral output). Accordingly, any variation in the average weighted productivity of an industry can be decomposed in a change in individual firm productivities for given output shares and in a reallocation of output shares for given individual firm productivities. Columns 11-13 of Tables 7 and 8 report the results of this decomposition averaged across years and sectors respectively. Weighted average productivity has grown at an average growth rate of 3 per cent a year, with industry reallocations driving 13 per cent of the overall increase. There is, however, a lot of variation across sectors and years. In particular, neither unweighted average productivity nor its components seem to show clear monotonic trends.

The comparison of productivity dispersion across industries may be sensitive to outliers within sectors and to differences in the scale of operations between them. We deal with the former issue by selecting the interquartile range as our benchmark measure of dispersion and address the latter through industry fixed effects.

5.3 Openness measure

To calculate $omega_{st}$ based on (9), we implement

$$\Omega_{st} = 1 + \tau_{st}^{-k_{st}} T_{st}^{-(\Lambda_{st}-1)} \tag{26}$$

where $\Lambda_{st} = k_{st}/(\sigma_s - 1)$ and $T_{st} = (f_X)_{st} / \left(\overline{f}_D\right)_s$ with $\overline{f}_D$ representing the domestic fixed cost averaged across time. For this we need the values of the elasticity of substitution (σ), the shape parameter of the Pareto distribution (k), the fixed cost ratio (T), as well as the transport cost parameter (τ).

The elasticity of substitution is derived as a by-product of the LPM estimation procedure. It is calculated as the reciprocal of the estimated coefficient of average deflated sales in equation (22). Estimates for each sector are reported in Table 9. The values of the estimated parameters, averaged across time, go from a minimum of 1.8 in sector 1 ('Food, beverages and tobacco') to a maximum of 5 in sector 7 ('Petroleum and coal'). Both the range of variation as well as the levels of this parameter look reasonable and

[13] Tables 7 and 8 also show that the TFP and thus the RMC distributions can be reasonably approximated by Pareto distributions. Formally, consider a random variable X (e.g., our TFP) with observed cumulative distribution $F(X)$. If the variable is distributed as a Pareto with shape parameter k_s, then the OLS estimate of the slope parameter in the regression of $\ln(1 - F(X))$ on $\ln(X)$ plus a constant is a consistent estimator of $-k_s$ and the corresponding R^2 is close to one. For all sectors the R^2 reported under k_s is far above 0.8.

comparable with the values reported in the literature. In particular, our estimates seem broadly in line with those in Broda and Weinstein (2006) who recover σ from the exact price index of several CES aggregate goods. They perform their estimations at a much finer level of disaggregation, actually obtaining tens of thousands of elasticities. However, at the SITC-3 level, which (though still finer than ours) is the closest to our disaggregation, they obtain a median value of 3.0, fairly in line with our own median value of 3.23. Nevertheless, their sectoral distribution is slightly skewed towards higher values when compared with ours (percentile 90: 7 against our 4.86; average: 5.9 against our 3.37; percentile 10: 1.9 against our 2.26). This again could stem from the fact that varieties should become increasingly substitutable as sectoral disaggregation gets finer.[14]

The shape parameter, which is sector specific, can be estimated in various ways. In our benchmark case, k is obtained by regressing the log of the rank of a firm within the TFP distribution on the log of the individual firm TFP (see Helpman et al. 2003). As reported in Tables 7 and 8, the goodness of fit of this regression is reasonably high in all sectors and years (R^2 is never below 0.8), suggesting that the ex-ante TFP distribution is fairly well approximated by a Pareto probability density function.

A proxy for the ratio of export to domestic fixed costs T can be obtained as follows. As suggested by Sutton (1991) and Syverson (2004), a measure of sunk entry costs in an industry is provided by the capital to value-added ratio multiplied by the market share of a median-sized firm. Since this can be seen as a measure of minimum efficient scale relative to the industry's total market size, the resulting product represents a proxy for the relative amount of capital required to achieve that scale. Since we are interested in $T = f_X/f_D$, we calculate that proxy separately for 'entrants in the export market' and 'non-exporters', then take the ratio of the former to the latter. Since only firms that start exporting in the current period and keep serving foreign markets for at least one period are counted as 'entrants' into the export market, such ratio accounts for the amount of 'extra-capital' needed in order to become an 'exporter'.[15] Moreover, f_D is made time-invariant by considering its average across years. This allows us to isolate the effect of differential fixed cost growth for exporters with respect to domestic producers in any given sector. All information needed is provided by the CADS dataset.

The measure of variable trade costs (τ) is obtained by estimating the gravity equation implied by our theoretical framework. Formally, as in Chaney (2006) and Helpman et al (2007), by using (2), (3), (6) and (8), export from the home to the foreign locations, $N_X \int_0^{C_X} \widetilde{P}_i \widetilde{D}_i dG(C_i)$, can be expressed as the product of parameters separately describing the locations' characteristics and an interaction term τ^{-k} that captures the dampening effects of variable trade costs.[16] In practice, the measure of variable trade costs between locations is obtained by assuming that τ depends on their bilateral distance

[14] At face value, our estimates would seem also compatible with those in Hummels (2001) who infers σ from the estimated parameter of the freight and tariff variable in a trade equation. The two sets of estimates, however, are not really consistent as the interpretation of the distance decay in trade equations differs between models with heterogeneous firms (our case) and models with representative firms (his case). See footnote 16 for further details.

[15] As stated in Section 4, throughout the analysis 'exporters' are firms exporting for more than half the years they are present in the sample.

[16] In trade models with representative firms, the interaction term is $\tau^{1-\sigma}$ instead of τ^{-k} so that the impact of variable trade costs on export flows depends on the elasticity substitution rather than on the skewness of the distribution of cost draws (see, e.g., Anderson and van Wincoop, 2003).

as well as on the existence of common border and language. Specifically, the variable trade cost between Italy and its trade partner h in sector s at time t is calculated as:

$$\left(\tau_{st}^{h}\right)^{-k_{st}} = (d^{h})^{-\delta_{st}} e^{(\nu_{st} Contig^{h} + \upsilon_{st} Lang^{h})} \tag{27}$$

where d^h is the bilateral distance, $Contig^h$ is a dummy that equals one if Italy and country h share a border, $Lang^h$ is a dummy variable that equals one if they share a common language.

The parameters δ_{st}, ν_{st} and υ_{st} are estimated from the following gravity equation

$$\ln(EXP_{st}^{l,h}) = EX_l + IM_h - \delta_{st}\ \ln(d^{l,h}) + \nu_{st}\ Contig^{l,h} + \upsilon_{st}\ Lang^{l,h} + \epsilon^{l,h} \tag{28}$$

where $EXP_{st}^{l,h}$ denotes export from country l to country h, EX_l and IM_h are origin and destination country fixed effects, and $\epsilon^{l,h}$ is a random error. Fixed effects pick up the influences stemming from exporting or importing country time-invariant characteristics.[17] The estimation is performed on a set of 100 countries (100 x 99 country pairs) selected on the basis of the total sectoral trade flow magnitude (exports + imports) over the period 1983-1999. Therefore, the set of countries is the same across years and within the same sector, but its composition may vary between sectors. Data on export flows come from the 'NBER-United Nations Trade Data' (Feenstra et al., 1997). Information on $Contig^{l,h}$, $Lang^{l,h}$ and $d^{l,h}$ comes from the 'dist-cepii' geographical dataset provided by the 'Centre d'Etude Prospectives et d'Informations Internationales' (CEPII). In particular, we measure $d^{l,h}$ as the simple (geodesic) distance calculated according to the great circle formula, which uses the latitudes and longitudes of the most important cities/agglomerations (in terms of population), and account also for internal distances.[18] In our benchmark specification, the value of ${\tau_{st}}^{-k_{st}}$ that appears in (26) is calculated as the simple average variable trade costs between Italy and its trade partners:

$${\tau_{st}}^{-k_{st}} = \sum_{h} (\tau_{st}^{h})^{-k_{st}} \tag{29}$$

with $(\tau_{st}^{h})^{-k_{st}}$ coming from (27).

A salient feature of the data is the high percentage of zeros in the bilateral trade matrix, which exceeds 60 per cent of the observations for several sector-year combinations. We deal with the resulting econometric issue following Santos Silva and Tenreyro (2006). Instead of the traditional log-linear version of (28), these authors estimate the model in its multiplicative form and propose a pseudo-maximum likelihood estimation (PMLE) technique, which has the additional advantage of addressing a serious heteroskedasticity problem induced by the log transformation.[19] Compared with the existing literature, our

[17] For similar approaches to gravity regressions, see Anderson and Van Wincoop (2003), Head and Mayer (2004), Redding and Venables (2004) in the case of representative firms as well as Chaney (2006), Helpman et al. (2007) in the case of heterogeneous firms.

[18] Following Head and Mayer (2002, 2004), the internal distance of location l is calculated as $(2/3)\sqrt{area^l/\pi}$, which models the average distance between a producer and a consumer on a stylized geography where all producers are centrally located and the consumers uniformly distributed across a disk-shaped surface.

[19] For a different methodology, see the two-stage estimation procedure proposed by Helpman et al. (2007). The approach by Santos Silva and Tenreyro (2006) has the advantage of avoiding the complex identification issues involved in the two-stage estimation.

estimated elasticities δ_{st}, which are reported by sector in Table 9 and by year in Table 10, take reasonable values. Their (unweighted) average across industries is 0.7, which is not very different from the one obtained in other comparable studies.[20] The ranking of sectors according to the estimated elasticities shows below-average levels for industries with high value added per unit-weight and for the so called 'Made in Italy' industries ('Jewelery and related articles, Musical instruments, Sport goods, Toys and Games', 'Electrical machinery and professional and scientific equipment', 'Transport equipment', 'Textiles', 'Leather and footwear').

The computed values of *omega* are reported in Tables 9 and 10 by sector and by year respectively. Interestingly enough, *omega* does not trend upwards monotonically and several periods are characterized by decreasing trade openness.

5.4 Results

The findings from our benchmark regression (25) are reported in the first row of Table 11. The univariate regression (Model [1]) clearly indicates that *cost dispersion and openness to trade are negatively correlated across industries.* The estimated parameters take negative values and are significantly different from zero. These findings are confirmed by adding industry and year fixed effects (Model [2]) and are also robust to assuming industry-heteroskedastic disturbances (Model [3]). As for the magnitude of the estimates, a two-standard-deviation decline in openness to trade is associated with a 0.9 per cent decrease of RMC dispersion, which is equivalent to a 7.6 per cent decline when evaluated in terms of the standard deviation of the explanatory variable. We also test Result 1 of the theoretical model, according to which more open sectors should exhibit lower values of central tendency. Table 11 shows that Result 1 is indeed supported by the data.

The negative impact of trade openness on the central tendency of the cost distribution is a common feature of the microeconometric trade literature. The negative impact on the dispersion is new and it is partially at odds with Syverson (2004) who shows that trade exposure has a positive (although not statistically significant) effect on productivity dispersion.[21] Our econometric results are, instead, largely consistent with those obtained by Helpman et al. (2003) and Bernard et al. (2006). The former show that a proxy for productivity dispersion is negatively correlated with the ratio of export sales to sales through Foreign Direct Investment (FDI) in the same industry. The latter find that low productivity plants are more likely to die in industries experiencing large trade costs reductions.

6 Robustness checks

This section checks to what extent our benchmark results on trade and dispersion are robust to various changes in the empirical specification.

[20]Santos Silva and Tenreyro (2006) obtain a value of 0.77 in their preferred specification, Helpman et al. (2007) obtain 0.802. When using PMLE, Del Gatto et al. (2006) find results that generate an average elasticity for Italy equal 1.166. The fact that they consider a smaller sample of highly integrated EU countries may explain the discrepancy.

[21]See Section 7 for possible explanations of this discrepancy.

6.1 Endogeneity

Despite the presence of industry and year fixed effects, regression (25) might suffer from problems of endogeneity and reverse causation.[22] To deal with this issue, we have first used lagged values of openness instead of its contemporaneous values. In this unreported analysis, the main findings are quite similar to our benchmark estimates. We have then instrumented the degree of openness to trade using Turkey's export share in world trade. Just like Italy, Turkey tends to be a net exporter in traditional sectors such as textiles and clothing, leather and footwear, etc. Thus, the ratio of Turkey's exports to worldwide trade flows should be correlated with Italy's openness to trade but reasonably uncorrelated with the productivity of its firms. Data on trade flows are drawn from the NBER-United Nations Trade Data (Feenstra et al., 1997). Although the hypothesis of inconsistent OLS estimates due to the presence of endogeneity can be rejected on the basis of the Hausman test, we nonetheless report the results of the Instrumental Variables analysis for all of our robustness checks. The corresponding estimates and standard errors are shown in the fourth column (Model [4]) of Tables 12, 13 and 14, together with other relevant statistics. These include the p-value of the F-test of the excluded instruments in the first stage regression, according to which Turkey's export share performs quite well as an instrument. Overall, the findings of the benchmark specification are confirmed.

6.2 Different dispersion and openness measures

Our results might depend on the specific measures we selected for dispersion, central tendency and trade openness. According to Tables 12, 13 and 14 that is not the case.

In particular, in Table 12 we consider three alternative measures of dispersion (90-10 Percentile Range, standard deviation, variance). Our benchmark results carry through. In Tables 13 and 14 we consider alternative measures of variable and fixed trade costs. First, we calculate the variable trade cost index as a weighted rather than a simple average:

$$\tau_{st}{}^{-k_{st}} = \sum_h z_{st}^h (\tau_{st}^h)^{-k_{st}} \qquad \text{with} \quad z_{st}^h = \frac{EXP_{st}^h}{\sum_h EXP_{st}^h} \tag{30}$$

where EXP_{st}^h is Italian exports to country h and $\sum_h EXP_{st}^h$ is total Italian exports. Through (27), this leads to a second openness measure $omega2$. Second, in our benchmark specification, we obtained a proxy for $T = f_X/f_D$ dividing the ratio of capital to value added for 'entrants in the export market' by the same ratio for 'nonexporters'. In order to focus on the effects of the changes in f_X, the ratio of capital to value added of non-esporters was made time-invariant by taking its average across years. By foregoing such averaging and implementing instead $T_{st} = (f_X)_{st}/(f_D)_{st}$, we get a third measure of openness $omega3$.

Tables 13 and 14 show that the negative impacts of openness on dispersion and central tendency survive. Table 14 also shows that measuring central tendency by the mean rather than by the median does not make much difference.

[22] This would be the case, for example, if the the political economy of trade policy depended on the dispersion of firm performance (Bombardini, 2005).

6.3 Generalized specification

So far we have heavily relied on the predictions of the theoretical model under the Pareto assumption. In the wake of Syverson (2004), we now try to assess the relation between RMC spreads and openness to trade within a more general framework. This is based on the following specification:

$$disp_{st} = \alpha + \beta open_{st} + \gamma sigma_s + \sum_t \iota_t I_t + \varepsilon_{st} \tag{31}$$

where *disp* is the dispersion measure and I_t denotes a vector of time fixed effects. Differently from the benchmark case, having removed the Pareto assumption and the industry fixed effects, we deal with the scale problem through standardized dispersion measures. They include the IQ range divided by the median ('St.IQ range') and the 90-10 percentile range divided by the median ('St.90-10 percentile range'). The regressions based on the former are presented in Tables 15 and 17, while those based on the latter are reported in Tables 16 and 18.

In (31) *sigma* is the industry-specific elasticity of substitution obtained through the LPM estimation procedure. It is included in the set of explanatory variables as Syverson (2004) shows that product differentiation is an important driver of productivity dispersion. As to trade opennes *open*, we use three alternative measures. The first is the 'propensity to export' (*pexp*), defined as the average export-to-output ratio for our sample of firms. The second is the index of variable transport costs τ used to calculate *omega* and it is thus an inverse measure of openness. Tables 15 and 16 list the regression results obtained with *pexp* and τ. They confirm the negative relation between trade openness and cost dispersion.

The third measure of openness is the export-to-output ratio computed from aggregate data. It is meant to remove possible problems of representativeness and sample bias affecting *pexp*. Data come from the Italian NIS (Coeweb statistics on external trade). These data also allow us to test the robustness of our results with respect to import intensity, which cannot be recovered from CADS. Tables 17 and 18 report the regression results obtained using for each industry the relevant ratio of: i) aggregate exports to gross value added (*AGGR-PEXP*), ii) aggregate imports to gross value added (*AGGR-PIMP*), iii) aggregate export-import flows to gross value added (*AGGR-PIMPEXP*). Benchmark results are again confirmed. In line with Syverson (2004), all specifications in Tables 15 to 18 also reveal a negative relation between product substitutability and cost dispersion.

7 Conclusion

Whether and how trade openness affects intra-industry firm heterogeneity may have important implications for the political economy of trade liberalization. In this respect, we have shown that a standard model with heterogeneous firms yields a key prediction: more open industries should feature smaller dispersion of firm marginal costs. We have shown that, when tested on Italian firm-level data, such prediction indeed finds empirical support. This has been achieved by estimating individual marginal costs through an innovative theory-based procedure due to Melitz (2000) that allows one to control not only for the standard transmission bias identified in firm-level TFP regressions but also for the omitted price bias due to imperfect competition.

Our findings are robust to alternative specifications and alternative measures of openness and dispersion. They are, however, at odds with those in Syverson (2004) who finds no significant effect of the propensity to export on productivity dispersion across US industries. The difference might stem from the fact that he measures productivity without correcting for the omitted price bias. His regressions, however, are based on a cross-section of 443 disaggregated industries (compared to our 18 more aggregated ones). This reduces the potential impact of the omitted price bias, which has been shown to be a decreasing function of product differentiation. Alternatively, the difference might be explained by the fact that he relies on a cross-section while we deal with a panel. More simply, the difference might mirror the bigger relevance that external trade has for Italy than for the US. As his and ours are the only existing studies that directly investigate the impact of trade openness on cost dispersion and they yield conflicting results, there is definitely a need for additional research on the topic.

References

[1] Anderson, James E., and Eric van Wincoop. 2003. "Gravity with Gravitas: A Solution to the Border Puzzle", *American Economic Review*, 93, 170-192.

[2] Aw, Bee Yan, Sukkyun Chung, and Mark J. Roberts. 2003. "Productivity, Output, and Failure: A Comparison of Taiwanese and Korean Manufacturers", *Economic Journal*, 113, 443-705.

[3] Baldwin, Richard E. 2005. "Heterogeneous Firms and Trade: Testable and Untestable Properties of the Melitz Model", NBER Working Paper No. 11471.

[4] Beck, Nathaniel, and Jonathan N. Katz. 1995. "What to (and not to) do with Time-Series Cross-section Data", *American Political Science Review*, 9, 634-647.

[5] Bernard, Andrew B., and J. Bradford Jensen. 1999. "Exceptional Exporter Performance: Cause, Effect, or Both?", *Journal of International Economics*, 47, 1-25.

[6] Bernard, Andrew B., Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum. 2003. "Plants and Productivity in International Trade", *American Economic Review*, 93, 1268-1290.

[7] Bernard, Andrew B., J. Bradford Jensen, and Peter K. Schott. 2006. "Trade costs, firms and productivity", *Journal of Monetary Economics*, 53, 917–937.

[8] Bombardini, Matilde. 2005. "Firm Heterogeneity and Lobby Participation", MIT, mimeograph.

[9] Broda, Christian, and David E. Weinstein. 2006. "Globalization and the Gains from Variety", *Quarterly Journal of Economics*, 121, 541-85.

[10] Chaney, Thomas. 2006. "Distorted Gravity: Heterogeneous Firms, Market Structure and the Geography of International Trade", University of Chicago, mimeograph.

[11] Cingano, Federico, and Fabiano Schivardi. 2004. "Identifying the Sources of Local Productivity Growth", *Journal of the European Economic Association*, 2, 720-742.

[12] Clerides, Sofronis K., Saul Lach, and James R. Tybout. 1998. "Is Learning By Exporting Important? Micro-Dynamic Evidence From Colombia, Mexico and Morocco", *Quarterly Journal of Economics*, 113, 903-947.

[13] Del Gatto, Massimo, Giordano Mion, and Gianmarco I.P. Ottaviano. 2006. "Trade Integration, Firm Selection and the Costs of Non-Europe", *CEPR Discussion Paper*, No. 5730.

[14] Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, Diego Puga, and Sebastien Roux. 2007. "The Productivity Advantage of Large Markets: Distinguishing Agglomeration from Firm Selection", University of Toronto, mimeograph.

[15] Falvey, Rod, David Greenaway, and Zhihong Yu (2004) "Intra-Industry Trade between Asymmetric Countries with Heterogeneous Firms", *GEP Research Paper*, No. 2004/05.

[16] Feenstra, Robert C., Robert E. Lipsey and Harry P. Bowen. 1997. "World Trade Flows, 1970-1992, with Production and Tariff Data", *NBER Working Paper*, No. 5910.

[17] Foster, Lucia, John Haltiwanger, and Chad Syverson. 2006. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?", mimeograph.

[18] Head, Keith, and Thierry Mayer. 2004. "Market Potential and the Location of Japanese Firms in the European Union", *Review of Economics and Statistics*, 86, 959-972.

[19] Helpman, Elhanan, Marc Melitz, and Stephen R. Yeaple. 2003. "Export versus FDI with Heterogeneous Firms", *American Economic Review*, 94(1), 300-316.

[20] Helpman, Elhanan, Marc Melitz, and Yona Rubinstein. 2007. "Trading Partners and Trading Volumes", *NBER Working Paper*, No. 12927.

[21] Hummels, David. 2001. "Towards a Geography of Trade costs", Purdue University, mimeograph.

[22] Klette, Tor J., and Zvi Griliches. 1996. "The Inconsistency of Common Scale Estimators when Output Prices are Unobserved and Endogenous", *Journal of Applied Econometrics*, 11, 343-361.

[23] Levinsohn, James, and Amil Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables", *Review of Economic Studies*, 70, 317-341.

[24] Mayer, Thierry, and Soledad Zignago. 2005. "Market Access in Global and Regional Trade", *CEPII Working Paper*, No. 02.

[25] Melitz, Marc. 2000. "Estimating Firm-Level Productivity in Differentiated Product Industry", Harvard University, mimeograph.

[26] Melitz, Marc. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity", *Econometrica*, 71, 1695-1725.

[27] Melitz, Marc, and Gianmarco I.P. Ottaviano. 2005. "Market size, Trade and Productivity", *NBER Working Paper*, No. 11393.

[28] Olley, Steve, and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry", *Econometrica*, 64, 1263-1297.

[29] Pavcnik, Nina. 2002. "Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants", *Review of Economic Studies*, 69, 245-276.

[30] Redding, Stephen, and Anthony J. Venables. 2004. "Economic Geography and International Inequality", *Journal of International Economics*, 62, 53-82.

[31] Roberts, Mark J., and James R. Tybout. 1997. "The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs", *American Economic Review*, 87, 545-564.

[32] Santos Silva, Joao, and Silvana Tenreyro. 2006. "The Log of Gravity", *Review of Economics and Statistics*, 88, 641-658.

[33] Sutton, John. 1991. *Sunk Costs and Market Structure*, Cambridge MA: MIT Press.

[34] Syverson, Chad. 2004. "Product substitutability and productivity dispersion", *Review of Economics and Statistics*, 86, 534-550.

Table 1: Sectoral disaggregation.

Code	**Short.**	**Description**
1	F	Food beverages and tobacco
2	TC	Textiles and Clothing
3	LF	Leather products and Footwear
4	W	Wood products except furniture
5	Pa	Paper products
6	PP	Printing and Publishing
7	PC	Petroleum and Coal
8	Ch	Chemicals
9	RP	Rubber and Plastic
10	NM	other Non-Metallic mineral products
11	Me	Metallic products
12	FM	Fabricated Metal products
13	Ma	Machinery except electrical
14	EM	Electrical Machinery and Professional and scientific equipment
15	Tr	Transport equipment
16	Fu	Furniture
17	JMST	Jewellery and related articles; Musical instruments; Sport goods; Toys and games
18	Oth.	Other manufacturing

Table 2: Sample characteristics: number of firms by year.

YEAR	Total	Exporters (%)
1983	8.871	34,8
1984	10.614	42,9
1985	11.579	42,5
1986	12.057	40,4
1987	12.102	40,3
1988	12.109	37,7
1989	12.552	40,2
1990	12.751	40,1
1991	12.769	38,9
1992	12.522	33,6
1993	12.133	31,1
1994	11.444	31,9
1995	8.435	32,2
1996	8.132	36,5
1997	8.018	39,0
1998	7.664	39,4
1999	6.926	37,1
Total	180.678	37,7

Table 3: Sample characteristics: number of firms by sector.

SECTOR	Total	Percentage	Total (%)	Exporters (%)
F	845	23.8%	7.9%	5.0%
TC	1,446	44.1%	13.6%	15.9%
LF	493	49.1%	4.6%	6.0%
W	259	23.1%	2.4%	1.5%
Pa	316	28.5%	3.0%	2.3%
PP	271	15.7%	2.5%	1.1%
PC	25	21.7%	0.2%	0.1%
Ch	650	35.8%	6.1%	5.8%
RP	637	38.5%	6.0%	6.1%
NM	760	30.7%	7.2%	5.8%
Me	446	34.3%	4.2%	3.8%
FM	1,121	34.3%	10.5%	9.6%
Ma	1,561	48.5%	14.7%	18.9%
EM	894	39.3%	8.4%	8.8%
Tr	278	40.3%	2.6%	2.8%
Fu	475	37.5%	4.5%	4.5%
JMST	107	51.8%	1.0%	1.4%
Oth.	45	48.1%	0.4%	0.5%
Total	10,628	37.7	100	100

Table 4: Sample characteristics: coverage 1991.

	Class of workers:												
SECTOR	1	2	3-5	6-9	10-15	16-19	20-49	50-99	100-199	200-249	250-499	500-999	1000-	Total
F	0.02	0.15	0.37	2.09	8.04	10.88	18.04	26.02	36.29	53.66	41.18	42.86	63.16	4.09
TC	0.04	0.11	0.87	2.54	5.95	5.94	14.42	31.99	40.64	48.24	52.87	58.33	125.00	6.12
LF	0.08	0.15	1.77	5.41	6.75	7.23	14.14	32.82	38.41	47.62	50.00	14.29	100.00	8.00
W	0.04	0.00	0.29	2.94	7.43	11.15	19.17	45.06	66.67	25.00	57.14	100.00	-	3.47
Pa	0.00	2.22	1.22	4.28	12.36	13.61	28.13	52.50	55.41	41.67	68.18	63.64	66.67	16.18
PP	0.00	0.00	0.24	1.06	3.71	6.62	16.00	35.32	42.42	42.11	28.57	66.67	66.67	3.43
PC	0.00	0.00	0.00	0.00	2.90	17.86	5.80	28.57	0.00	50.00	-	62.50	50.00	6.40
Ch	0.00	0.95	2.60	8.81	15.96	18.18	38.11	52.65	45.64	43.33	47.31	71.43	62.79	19.17
RP	0.00	0.51	0.66	2.48	8.28	10.85	26.43	48.99	53.44	52.63	75.00	93.33	66.67	12.56
NM	0.00	0.48	1.51	2.81	7.97	11.92	22.88	42.21	50.76	62.07	66.04	79.17	85.71	10.20
Me	1.02	1.39	3.85	5.83	15.14	20.71	30.64	50.69	51.52	60.00	57.45	72.73	50.00	25.33
FM	0.00	0.00	0.69	1.24	4.95	5.96	18.20	37.17	42.46	67.44	49.23	44.44	0.00	6.15
Ma	0.10	0.36	1.13	3.38	9.69	10.68	25.21	49.00	56.08	45.00	52.26	71.43	69.70	13.49
EM	0.02	0.11	0.56	2.29	8.28	12.06	21.76	48.41	59.38	47.62	60.19	50.00	82.61	6.48
Tr	0.00	0.00	1.04	2.78	5.54	5.10	16.57	30.59	36.51	40.91	55.00	41.94	90.00	10.87
Fu	0.09	0.11	0.85	1.74	5.84	8.90	20.17	44.55	53.77	52.94	60.00	40.00	0.00	7.22
JMST	0.00	0.00	0.72	0.62	5.79	3.49	21.00	35.53	50.00	0.00	100.00	100.00	-	4.21
Oth.	0.00	0.00	0.25	0.91	2.40	10.64	13.38	25.58	28.57	0.00	40.00	0.00	-	2.44
Total	0.03	0.15	0.75	2.53	7.27	9.17	20.05	40.10	47.20	49.90	53.35	59.24	71.56	7.41

Source: ISTAT and CADS.

Table 5: Sample characteristics: coverage 1996.

	Class of workers:													
SECTOR	1	2	3-5	6-9	10-15	16-19	20-49	50-99	100-199	200-249	250-499	500-999	1000-	Total
F	0.00	0.00	0.00	0.60	4.13	8.80	19.87	42.86	42.98	50.94	49.23	51.72	87.50	5.20
TC	0.00	0.00	0.19	0.97	1.53	3.92	8.64	31.91	45.87	62.16	53.68	75.00	100.00	8.00
LF	0.00	0.00	-	0.67	2.59	2.29	8.24	34.89	48.46	66.67	68.18	12.50	100.00	7.41
W	0.00	0.00	0.00	-	3.05	1.86	8.63	36.62	52.08	25.00	50.00	100.00	-	5.43
Pa	0.00	0.00	-	-	1.43	8.81	16.94	59.59	58.33	56.25	75.00	100.00	33.33	15.69
PP	0.00	0.00	-	-	0.13	1.06	6.46	33.15	49.41	46.67	26.92	38.46	33.33	2.00
PC	0.00	0.00	0.00	0.00	3.13	3.33	18.75	23.53	18.18	100.00	40.00	33.33	-	6.51
Ch	0.00	0.00	0.65	2.25	3.72	7.34	26.43	54.39	52.54	61.76	54.43	60.87	66.67	16.86
RP	0.00	0.00	0.00	-	0.74	2.97	14.00	42.06	53.99	72.73	57.89	100.00	100.00	11.09
NM	0.00	0.16	0.19	0.67	0.97	2.49	11.35	35.17	52.76	52.78	79.59	76.19	60.00	6.78
Me	0.00	0.00	-	2.78	4.08	8.96	18.33	56.04	55.73	58.82	59.52	36.36	46.15	19.15
FM	0.00	0.00	0.00	0.53	0.22	0.87	6.03	38.90	50.34	55.81	64.06	66.67	0.00	5.07
Ma	0.00	0.00	0.07	0.48	0.31	1.82	10.25	43.41	53.31	47.87	58.90	65.00	54.55	10.49
EM	0.00	0.00	0.00	0.58	1.17	3.14	9.00	37.29	45.77	80.56	60.75	50.00	84.38	6.62
Tr	0.00	0.00	0.00	0.85	1.44	2.67	9.14	31.46	43.06	56.52	65.85	45.45	64.29	10.04
Fu	-	-	-	0.11	0.23	0.91	8.83	47.85	56.38	56.25	59.26	80.00	0.00	9.04
JMST	0.00	0.00	0.00	0.67	0.41	0.90	12.66	34.92	56.00	50.00	66.67	200.00	-	4.28
Oth.	0.00	0.00	0.00	0.00	1.52	0.00	5.34	28.57	31.03	0.00	40.00	0.00	100.00	2.01
Total	0.00	0.01	0.08	0.69	1.36	3.04	10.59	39.54	49.60	57.20	58.28	60.50	67.76	7.24

Source: ISTAT and CADS.

Table 6: Descriptive statistics for key variables.

SECTOR	TOTAL					EXPORTERS				
	Value Added*	Revenues*	Exports*	Capital*	Employees	Value Added*	Revenues*	Exports*	Capital*	Employees
All Sectors - 1991	4.918	16.521	2.407	11.000	107	6.832	14.928	22.621	6.183	148
of which:										
F	6,066	27,094	1,242	14,552	112	10,843	28,641	44,968	5,758	192
TC	3,280	11,586	2,028	7,361	92	4,236	8,919	15,229	4,361	117
LF	1,916	8,470	2,391	2,696	55	2,268	3,341	9,814	4,355	65
W	1,870	6,987	404	5,181	49	2,787	7,642	9,983	1,793	75
Pa	5,282	17,648	1,827	15,960	99	10,179	31,414	32,612	6,369	179
PP	5,436	15,272	450	9,011	100	5,697	14,621	16,671	3,250	116
PC	34,269	157,659	2,164	138,286	474	24,223	33,553	85,364	9,379	366
Ch	9,516	36,435	4,001	24,327	169	14,351	42,028	54,352	11,190	260
RP	4,434	13,474	2,166	11,608	97	6,951	19,393	20,345	5,483	151
NM	4,305	11,434	1,397	15,297	93	4,969	15,056	13,432	4,470	114
Me	7,764	29,302	3,885	23,707	166	17,004	45,887	62,029	13,035	359
FM	2,543	7,876	1,052	5,457	60	3,576	7,938	10,934	3,010	81
Ma	4,463	14,351	3,822	7,359	106	5,671	9,249	18,736	7,430	133
EM	9,337	25,036	4,404	14,108	187	14,410	22,472	39,456	10,684	271
Tr	7,027	21,725	2,741	14,259	184	8,345	16,508	25,130	6,583	226
Fu	2,089	7,698	1,262	3,933	61	2,559	4,756	9,980	3,093	72
JMST	2,177	8,602	2,383	4,201	59	2,324	4,617	8,737	3,777	61
Oth.	2,137	6,601	1,070	4,296	61	2,167	4,812	6,070	2,016	67
All Sectors - 1983	4.680	15.034	1.419	15.782	150	6.892	24.285	22.100	4.058	224
All Sectors - 1999	8.669	33.922	13.393	18.988	155	12.573	27.671	48.104	19.064	217

Note: *Thousands of 1995 euros.

Table 7: Productivity and productivity dispersion by sector: LPM estimation.

	Productivity ($\hat{a}^{LPM}$)		IQ Range		90-10 Range		k_s			Decomposition of Productivity Growth		
SECTOR	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.	R^2	Aggr. Pr.	Unweighted Pr.	Covariance
F	0.004	0.001	1.040	0.078	2.365	0.164	1.120	0.059	0.830	0.030	0.029	0.001
TC	0.518	0.067	0.832	0.047	1.883	0.158	1.466	0.059	0.876	0.022	0.017	0.006
LF	3.736	0.527	0.599	0.044	1.257	0.058	1.867	0.098	0.831	0.005	0.003	0.001
W	4.949	0.634	0.550	0.042	1.179	0.077	2.047	0.086	0.831	0.007	0.007	0.000
Pa	0.173	0.031	0.751	0.082	1.619	0.136	1.453	0.092	0.798	0.040	0.042	-0.002
PP	0.754	0.115	0.728	0.083	1.488	0.146	1.701	0.072	0.862	0.021	0.024	-0.003
PC	3.668	1.399	0.670	0.140	1.413	0.499	1.549	0.171	0.837	-0.007	0.006	-0.013
Ch	0.078	0.016	0.847	0.051	1.818	0.106	1.388	0.091	0.828	0.058	0.045	0.013
RP	0.458	0.053	0.612	0.032	1.359	0.047	1.787	0.078	0.840	0.016	0.012	0.004
NM	0.096	0.017	0.757	0.050	1.673	0.072	1.542	0.059	0.871	0.037	0.031	0.006
Me	0.197	0.042	0.665	0.044	1.480	0.122	1.694	0.088	0.856	0.062	0.045	0.017
FM	0.343	0.059	0.637	0.042	1.369	0.064	1.766	0.121	0.845	0.062	0.060	0.002
Ma	1.556	0.212	0.588	0.023	1.232	0.053	1.949	0.116	0.833	0.013	0.010	0.002
EM	0.639	0.137	0.691	0.038	1.492	0.056	1.688	0.101	0.848	0.050	0.050	0.000
Tr	0.706	0.091	0.636	0.049	1.387	0.112	1.788	0.113	0.851	0.006	0.006	0.000
Fu	2.294	0.280	0.492	0.040	1.051	0.060	2.145	0.145	0.844	0.008	0.005	0.004
JMST	3.244	0.464	0.552	0.062	1.316	0.156	1.944	0.097	0.911	0.004	0.002	0.002
Oth.	1.444	0.133	0.704	0.138	1.401	0.258	1.660	0.150	0.869	-0.027	-0.031	0.004
TOTAL	1.381	1.546	0.686	0.143	1.488	0.338	1.697	0.266	0.848	0.030	0.027	0.004

Note: Data are averaged across years.

Table 8: Productivity and productivity dispersion by year: LPM estimation.

	Productivity ($\hat{a}^{LPM}$)		IQ Range		90-10 Range		k_s			Decomposition of Productivity Growth		
YEAR	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.	$\bar{R}^2$	Aggr. Pr.	Unweighted Pr.	Covariance
1983	0.999	1.179	0.624	0.133	1.442	0.322	1.773	0.264	0.872	-	-	-
1984	1.026	1.183	0.652	0.138	1.543	0.487	1.756	0.276	0.865	0.077	0.053	0.023
1985	1.086	1.233	0.669	0.144	1.473	0.349	1.706	0.225	0.842	0.082	0.076	0.006
1986	1.174	1.273	0.680	0.149	1.474	0.377	1.750	0.257	0.849	0.095	0.097	-0.002
1987	1.218	1.292	0.689	0.150	1.481	0.291	1.718	0.275	0.845	0.087	0.068	0.020
1988	1.269	1.364	0.682	0.138	1.473	0.318	1.704	0.269	0.842	0.037	0.019	0.018
1989	1.292	1.399	0.675	0.117	1.424	0.256	1.714	0.278	0.832	0.035	0.036	-0.001
1990	1.352	1.464	0.680	0.124	1.437	0.317	1.667	0.265	0.827	0.056	0.049	0.007
1991	1.415	1.538	0.697	0.155	1.468	0.307	1.651	0.275	0.830	0.030	0.003	0.027
1992	1.505	1.673	0.709	0.133	1.502	0.295	1.598	0.234	0.821	0.019	0.041	-0.022
1993	1.515	1.662	0.733	0.158	1.540	0.370	1.565	0.244	0.823	0.017	0.022	-0.005
1994	1.654	1.857	0.688	0.140	1.568	0.294	1.623	0.232	0.848	0.040	0.042	-0.001
1995	1.684	1.962	0.703	0.156	1.553	0.396	1.693	0.273	0.872	-0.080	-0.055	-0.024
1996	1.603	1.857	0.690	0.149	1.473	0.340	1.754	0.284	0.867	-0.037	-0.025	-0.012
1997	1.557	1.764	0.670	0.142	1.460	0.348	1.766	0.288	0.868	-0.009	-0.026	0.017
1998	1.569	1.840	0.707	0.152	1.486	0.328	1.727	0.303	0.856	-0.043	-0.043	-0.001
1999	1.558	1.823	0.714	0.168	1.497	0.394	1.690	0.281	0.852	-0.017	-0.017	0.000
Total	1.381	1.546	0.686	0.143	1.488	0.338	1.697	0.266	0.848	0.030	0.027	0.004

Note: Data are averaged across industries.

Table 9: Descriptive statistics for regressors by sector.

	OMEGA		δ_s		σ_s
SECTOR	Mean	Std.Dev.	Mean	Std.Dev.	Mean
F	1.781	0.275	-0.716	0.048	1.835
TC	2.807	1.400	-0.678	0.098	2.882
LF	5.592	2.895	-0.504	0.066	4.086
W	1.385	0.296	-0.903	0.059	4.367
Pa	1.182	0.074	-0.912	0.031	2.258
PP	9.396	2.270	-0.653	0.012	3.303
PC	1.018	0.015	-1.271	0.079	5.055
Ch	1.297	0.076	-0.813	0.021	2.391
RP	1.580	0.127	-0.734	0.019	3.165
NM	2.431	0.707	-0.624	0.048	2.969
Me	1.497	0.231	-0.785	0.055	2.913
FM	2.996	0.551	-0.681	0.022	2.552
Ma	2.868	0.487	-0.571	0.045	3.454
EM	6.774	2.194	-0.478	0.025	3.057
Tr	4.143	2.674	-0.595	0.080	3.995
Fu	4.227	6.808	-0.765	0.106	3.856
JMST	6.809	3.240	-0.430	0.054	4.857
Oth.	4.215	2.052	-0.588	0.083	3.624
TOTAL	3.489	3.185	-0.706	0.198	3.368

Note: Data are averaged across years.

Table 10: Descriptive statistics for regressors by year.

	OMEGA		δ_s	
YEAR	Mean	Std.Dev.	Mean	Std.Dev.
1983	3.762	3.365	-0.682	0.179
1984	3.585	2.470	-0.689	0.175
1985	4.106	4.951	-0.676	0.182
1986	4.516	5.191	-0.672	0.177
1987	4.104	3.737	-0.691	0.193
1988	4.515	3.501	-0.659	0.197
1989	4.670	3.930	-0.665	0.209
1990	4.238	3.200	-0.673	0.223
1991	3.020	2.450	-0.707	0.213
1992	3.251	3.001	-0.725	0.214
1993	3.000	2.605	-0.715	0.210
1994	2.579	1.646	-0.733	0.212
1995	2.876	2.640	-0.738	0.216
1996	2.846	2.416	-0.742	0.215
1997	2.923	2.803	-0.743	0.214
1998	2.510	1.676	-0.747	0.193
1999	2.818	2.157	-0.741	0.195
Total	3.489	3.185	-0.706	0.198

Note: Data are averaged across industries.

Table 11: Benchmark results.

Dependent Variable	Regressors and statistics	Model [1]	Model [2]	Model [3]
RMC Dispersion	omega	-0.794***	-0.038***	-0.038**
		(0.126)	(0.014)	(0.015)
	N	279	279	279
	R^2	0.08	-	-
RMC Central Tendency	omega	-0.719***	-0.024*	-0.024***
		(0.117)	(0.012)	(0.008)
	N	279	279	279
	R^2	0.08	-	-
Industry fixed-effects		no	yes	yes
Time fixed-effects		no	yes	yes

Notes: Key to specifications by column: [1] Univariate OLS regression; [2] Fixed-effects regression; [3] Prais-Winsten with panel corrected standard errors (PCSE) regression. Robust standard errors (Model [1] and Model [2]) and panel-corrected standard errors (Model [3]) in parenthesis.

Dispersion measure: IQ Range. Central tendency measure: median. All variables in logs.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 12: Robustness checks: RMC regressions with different dispersion measures.

Dependent Variable	Regressors and statistics	Model [1]	Model [2]	Model [3]	Model [4]
IQ Range	omega	-0.794***	-0.038***	-0.038**	-0.224***
		(0.126)	(0.014)	(0.015)	(0.080)
	N	279	279	279	279
	R^2	0.08	-	-	-
	$F > 0$ on excluded instr.	-	-	-	0.000
90-10 Range	omega	-0.810***	-0.017	-0.017	-0.246***
		(0.127)	(0.016)	(0.015)	(0.075)
	N	279	279	279	279
	R^2	0.08	-	-	-
	$F > 0$ on excluded instr.	-	-	-	0.000
Std.Dev.	omega	-0.803***	-0.030**	-0.030**	-0.249***
		(0.127)	(0.014)	(0.013)	(0.080)
	N	279	279	279	279
	R^2	0.08	-	-	-
	$F > 0$ on excluded instr.	-	-	-	0.000
Variance	omega	-1.607***	-0.060**	-0.060**	-0.498***
		(0.253)	(0.027)	(0.027)	(0.161)
	N	279	279	279	279
	R^2	0.08	-	-	-
	$F > 0$ on excluded instr.	-	-	-	0.000
Industry fixed-effects		no	yes	yes	yes
Time fixed-effects		no	yes	yes	yes

Notes: Key to specifications by column: [1] Univariate OLS regression; [2] Fixed-effects regression; [3] Prais-Winsten with panel corrected standard errors (PCSE) regression; [4] IV regression (instrumented: omega; instrument: export quote of Turkey). Robust standard errors (Model [1] and Model [2]) and panel-corrected standard errors (Model [3] and Model [4]) in parenthesis. All variables in logs. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 13: Robustness checks: RMC dispersion regressions with different measures of trade openness.

Dependent Variable	Regressors and statistics	Model [1]	Model [2]	Model [3]	Model [4]
IQ Range	omega2	-11.809***	-0.390**	-0.390*	-6.109**
		(1.981)	(0.169)	(0.234)	(2.520)
	N	279	279	279	279
	R^2	0.05	-	-	-
	$F > 0$ on excluded instr.	-	-	-	0.049
IQ Range	omega3	-0.763***	-0.037**	-0.037***	-0.270***
		(0.107)	(0.015)	(0.012)	(0.080)
	N	289	289	289	279
	R^2	0.11	-	-	-
	$F > 0$ on excluded instr.	-	-	-	0.000
Industry fixed-effects		no	yes	yes	yes
Time fixed-effects		no	yes	yes	yes

Notes: Key to specifications by column: [1] Univariate OLS regression; [2] Fixed-effects regression; [3] Prais-Winsten with panel corrected standard errors (PCSE) regression; [4] IV regression (instrumented: omega; instrument: export quote of Turkey). Robust standard errors (Model [1] and Model [2]) and panel-corrected standard errors (Model [3] and Model [4]) in parenthesis. All variables in logs. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 14: Robustness checks: RMC central tendency regressions.

Dependent Variable	Regressors and statistics	Model [1]	Model [2]	Model [3]	Model [4]
Median	omega	-0.719***	-0.024*	-0.024***	-0.228***
		(0.117)	(0.012)	(0.008)	(0.064)
	N	279	279	279	279
	R^2	0.08	-	-	-
	$F > 0$ on excluded instr.	-	-	-	0.000
Mean	omega	-0.732***	-0.017	-0.017**	-0.225***
		(0.118)	(0.012)	(0.008)	(0.064)
	N	279	279	279	279
	R^2	0.08	-	-	-
	$F > 0$ on excluded instr.	-	-	-	0.000
Median	omega2	-10.778***	-0.214	-0.214*	-6.140***
		(1.832)	(0.141)	(0.111)	(2.152)
	N	279	279	279	279
	R^2	0.05	-	-	-
	$F > 0$ on excluded instr.	-	-	-	0.049
Median	omega3	-0.690***	-0.030**	-0.030**	-0.259***
		(0.100)	(0.014)	(0.012)	(0.070)
	N	289	289	289	289
	R^2	0.11	-	-	-
	$F > 0$ on excluded instr.	-	-	-	0.000
Industry fixed-effects		no	yes	yes	yes
Time fixed-effects		no	yes	yes	yes

Notes: Key to specifications by column: [1] Univariate OLS regression; [2] Fixed-effects regression; [3] Prais-Winsten with panel corrected standard errors (PCSE) regression; [4] IV regression (instrumented: omega; instrument: export quote of Turkey). Robust standard errors (Model [1] and Model [2]) and panel-corrected standard errors (Model [3] and Model [4]) in parenthesis. All variables in logs. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 15: Robustness checks: Generalized specification.

Dependent Variable	Regressor and statistics	Model [1]	Model [2]	Model [3]
IQ Range	pexp	-0.267***	-0.308***	-0.308***
		(0.070)	(0.079)	(0.024)
	σ	-1.896***	-1.893***	-1.893***
		(0.092)	(0.094)	(0.015)
	N	289	289	289
	R^2	0.74	-	-
IQ Range	τ	0.438**	0.493**	0.493***
		(0.206)	(0.213)	(0.053)
	σ	-1.915***	-1.914***	-1.914***
		(0.095)	(0.097)	(0.018)
	N	289	289	289
	R^2	0.73	-	-
Industry fixed-effects		no	yes	yes
Time fixed-effects		no	yes	yes

Notes: Key to specifications by column: [1] Univariate OLS regression; [2] Fixed-effects regression; [3] Prais-Winsten with panel corrected standard errors (PCSE) regression; [4] IV regression (instrumented: omega; instrument: export quote of NAC). Robust standard errors (Model [1] and Model [2]) and panel-corrected standard errors (Model [3] and Model [4]) in parenthesis. All variables in logs.
* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 16: Robustness checks: Generalized specification.

Dependent Variable	Regressor and statistics	Model [1]	Model [2]	Model [3]
90-10 Range	pexp	-0.271***	-0.299***	-0.299***
		(0.072)	(0.082)	(0.025)
	σ	-1.912***	-1.909***	-1.909***
		(0.094)	(0.096)	(0.016)
	N	289	289	289
	R^2	0.74	-	-
90-10 Range	τ	0.449**	0.509**	0.509***
		(0.208)	(0.214)	(0.060)
	σ	-1.931***	-1.930***	-1.930***
		(0.097)	(0.099)	(0.017)
	N	289	289	289
	R^2	0.73	-	-
Industry fixed-effects		no	yes	yes
Time fixed-effects		no	yes	yes

Notes: Key to specifications by column: [1] Univariate OLS regression; [2] Fixed-effects regression; [3] Prais-Winsten with panel corrected standard errors (PCSE) regression; [4] IV regression (instrumented: omega; instrument: export quote of NAC). Robust standard errors (Model [1] and Model [2]) and panel-corrected standard errors (Model [3] and Model [4]) in parenthesis. All variables in logs.
* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 17: Robustness checks: Generalized specification.

Dependent Variable	Regressor and statistics	Model [1]		Model [2]		Model [3]	
St.IQ Range	AGGR-PEXP	-572.685***	-488.323***	-231.532	-497.935***	-231.532*	-497.935***
		(188.411)	(168.735)	(201.623)	(174.529)	(136.615)	(12.036)
	σ	-	-49.728***	-	-49.685***	-	-49.685***
		-	(8.619)	-	(8.86)	-	(1.478)
	N	128	128	128	128	128	128
	R^2	0.06	0.25	-	-	-	-
St.IQ Range	AGGR-PIMP	-404.095*	-773.537***	-261.403	-781.324***	-261.403***	-
		(221.906)	(195.717)	(207.546)	(201.770)	(92.759)	-
	σ	-	-61.091***	-	-61.184***	-	-
		-	(8.68)	-	(8.925)	-	-
	N	128	128	128	128	128	-
	R^2	0.02	0.29	-	-	-	-
St.IQ Range	AGGR-PIMPEXP	-342.405***	-403.749***	-189.556	-410.416***	-189.556***	-410.416***
		(118.066)	(102.772)	(126.465)	(106.165)	(69.528)	(8.980)
	σ	-	-54.905***	-	-54.955***	-	-54.955***
		-	(8.403)	-	(8.634)	-	(1.564)
	N	128	128	128	128	128	128
	R^2	0.06	0.29	-	-	-	-
Industry fixed-effects		no	no	yes	no	yes	no
Time fixed-effects		no	no	yes	yes	yes	yes

Notes: Key to specifications by column: [1] Univariate OLS regression; [2] Fixed-effects regression; [3] Prais-Winsten with panel corrected standard errors (PCSE) regression; [4] IV regression (instrumented: omega; instrument: export quote of NAC). Robust standard errors (Model [1] and Model [2]) and panel-corrected standard errors (Model [3] and Model [4]) in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 18: Robustness checks: Generalized specification.

Dependent Variable	Regressor and statistics	Model [1]		Model [2]		Model [3]	
St. 90-10 Percentile range	AGGR-PEXP	-1291.522***	-1101.856***	-135.652	-1116.801***	-135.652	-1116.801***
		(423.977)	(379.790)	(369.603)	(393.051)	(303.131)	(37.246)
	σ	-	-111.801***	-	-111.735***	-	-111.735***
		-	(19.400)	-	(19.954)	-	(2.450)
	N	128	128	128	128	128	128
	R^2	0.06	0.25	-	-	-	-
St. 90-10 Percentile range	AGGR-PIMP	-912.239*	-1743.241***	-6.567	-1753.278***	-6.567	-1753.278***
		(499.383)	(440.532)	(381.189)	(454.472)	(205.748)	(60.428)
	σ	-	-137.415***	-	-137.535***	-	-137.535***
		-	(19.538)	-	(20.102)	-	(3.173)
	N	128	128	128	128	128	128
	R^2	0.02	0.29	-	-	-	-
St. 90-10 Percentile range	AGGR-PIMPEXP	-772.465***	-910.424***	-56.293	-920.748***	-56.293	-920.748***
		(265.679)	(231.309)	(232.933)	(239.135)	(173.659)	(28.080)
	σ	-	-123.478***	-	-123.555***	-	-123.555***
		-	(18.913)	-	(19.448)	-	(2.694)
	N	128	128	128	128	128	128
	R^2	0.06	0.29	-	-	-	-
Industry fixed-effects		no	no	yes	no	yes	no
Time fixed-effects		no	no	yes	yes	yes	yes

Notes: Key to specifications by column: [1] Univariate OLS regression; [2] Fixed-effects regression; [3] Prais-Winsten with panel corrected standard errors (PCSE) regression; [4] IV regression (instrumented: omega; instrument: export quote of NAC). Robust standard errors (Model [1] and Model [2]) and panel-corrected standard errors (Model [3] and Model [4]) in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%.