Gravity with Granularity*

Holger Breinlich† Harald Fadinger‡ Volker Nocke§ Nicolas Schutz¶

September 12, 2024

Abstract

We demonstrate that the estimation of gravity equations of trade flows suffers from an omitted variable bias when firms are granular and behave oligopolistically. We show how to correct for this bias in the estimation of both firm- and industry-level gravity. Using French and Chinese export data, we find that the oligopoly bias leads to a substantial underestimation of the effects of distance on trade flows. In a calibrated version of the model, the welfare gains from a trade liberalization are found to be almost twice as large under oligopoly as under monopolistic competition.

Keywords: Gravity Equation, Oligopoly, Trade Liberalization, Trade Elasticity

Journal of Economic Literature Classification: F12, F14, L13

---

*We thank Thibault Fally, Rob Feenstra, Keith Head, Kadee Russ, Stephen Redding, Jon Vogel, Yongjin Wang, seminar participants at Berkeley, DICE, Essex, Geneva, LSE, Munich, Nanjing, Nottingham, PSE, Stockholm School of Economics, and Surrey, and conference participants at Bayreuth Trade and IO Workshop 2022, the 2023 Christmas Meeting of German Economists Abroad, EARIE 2019, ENTER Jamboree 2022, ERWIT-CURE 2023, ETSG 2019, NBER Summer Institute 2023, SED 2023, the 2019 meeting of the international economics committee of the *Verein für Socialpolitik*, and WIEN 2022, for helpful suggestions. We gratefully acknowledge financial support from the German Research Foundation (DFG) through CRC TR 224 (Projects B03 and B06). Access to French confidential data, on which some of this work is based, has been made possible within a secure environment offered by CASD – Centre d'accès sécurisé aux données (Ref. 10.34724/CASD).

†University of Surrey, CEP and CEPR. Email: h.breinlich@surrey.ac.uk.

‡University of Vienna, CEPR and MaCCI. Email: harald.fadinger@univie.ac.at

§Department of Economics and MaCCI, University of Mannheim. Also affiliated with CEPR. Email: volker.nocke@gmail.com.

¶Department of Economics and MaCCI, University of Mannheim. Also affiliated with CEPR. Email: schutz@uni-mannheim.de.

1 Introduction

Gravity equations have been the predominant tool for analyzing the determinants of bilateral trade flows since their introduction by Tinbergen (1962) over 60 years ago. In their most basic form, gravity equations predict that trade between countries is a log-linear function of the economic mass of the two trading partners and bilateral frictions such as distance or tariffs. Even in this simple form, gravity equations have substantial explanatory power, often explaining in excess of 70-80% of the variation in the trade flows between countries. Starting with Anderson (1979), researchers have shown that gravity equations can be derived from a number of mainstream theoretical frameworks, allowing a tight link to economic welfare analysis. Not surprisingly then, gravity equations have become the workhorse tool for evaluating trade-related economic policies, such as tariffs, trade agreements or WTO membership.

Despite the rapid progress that research on gravity equations has made over the past decades, existing approaches remain at odds with a key stylized fact about international trade, however: much of world trade is dominated by a small number of large firms. The classic example is the market for wide-bodied passenger aircraft which comprises just two firms (Airbus and Boeing); but the markets of many other tradable goods such as cars, mobile phones or television sets are also dominated by a handful of large producers. That is, in the language of Gaubert and Itskhoki (2021), trade flows are "granular". Given their size, it seems likely that such "granular" firms enjoy substantial market power and have incentives to internalize the effects of their actions on aggregate market outcomes. In this paper, we evaluate the consequences of oligopolistic behavior for the estimation of gravity equations.

Under oligopoly, standard approaches to gravity estimation deliver inconsistent estimates of key parameters, such as the trade elasticity with respect to distance. The reason is that markups co-vary systematically with bilateral variable trade costs (e.g., distance or tariffs) and are contained in the error term of the gravity equation. The key intuition is that firms selling in destinations with higher bilateral trade costs face higher marginal costs, and that these higher marginal costs are incompletely passed through under oligopoly.[1] This induces a classical omitted variable bias, which leads to an under-estimation of the trade elasticity: The value of exports does not fall as much with variable trade costs as it would fall if markups were held constant, because firms systematically reduce markups when selling to destinations with higher trade costs.

[1]For evidence on incomplete cost pass-through in the industrial organization and international trade literatures, see Feenstra (1989), Nakamura and Zerom (2010), Burstein and Gopinath (2014), Ganapati, Shapiro, and Walker (2020), and Genakos and Pagliero (2022).

We derive firm- and industry-level gravity equations from a rich heterogeneous-firm model with oligopolistic competition and product differentiation based on Atkeson and Burstein (2008). Consumers have CES preferences with industry-specific demand elasticities and firms draw idiosyncratic productivity and quality shocks. We show that this model generates a gravity equation both at the firm and industry level. We also propose methods for consistent estimation of gravity-equation parameters.

Specifically, we show how to eliminate the oligopoly bias by constructing a correction term that purges observed trade flows from oligopolistic market power effects. At the firm level, the correction term uses information on firms' market shares and demand elasticities. At the industry level, the correction term takes the form of an origin-destination-level Herfindahl index (HHI) of exporters multiplied by the exporting country's aggregate market share in the destination market. This is intuitive: exporters' markups are high if exports are concentrated in a small number of firms that have a large aggregate market share in the destination.

In our empirical applications, we use firm- and industry-level data on exports of French and Chinese firms to European countries, and therefore focus on distance as the only bilateral trade cost variable.[2] We show that failing to account for oligopoly leads to a substantial underestimation of the distance elasticity of trade flows.[3] At the firm level, the average oligopoly bias is in excess of 40%. At the industry level, the bias is around 10% for the average industry but it is substantially larger in a significant minority of industries, in which exports tend to be highly concentrated.

To confirm the validity of our empirical approach, we perform a detailed Monte Carlo study. We calibrate our rich heterogeneous-firm CES oligopoly model to match key statistics of the French and Chinese micro-level trade data, and use it to generate a simulated dataset. We then run firm- and industry-level gravity regressions on that dataset, and find that our oligopoly corrections do very well in recovering the distance coefficient. By contrast, without the oligopoly correction, we obtain a bias of similar magnitude to that in the regressions run on the actual data.

Finally, we use our calibrated model to evaluate the welfare effects of a 10% trade cost reduction. We find that the resulting welfare gains are almost twice as high under oligopoly as under monopolistic competition. This is driven both by the larger estimated distance coefficient under oligopoly compared to monopolistic competition, and by additional pro-

[2]Because of the restriction to European destinations (for reasons outlined below), there is insufficient variation to include other common gravity variables. However, our methodology naturally applies also to policy-relevant variables such as tariffs and regional trade agreements.

[3]As a result, the estimated distance elasticity of trade costs is biased downwards. Using price data for shipments and estimates of cost pass-through to account for variable markups, Atkin and Donaldson (2015) also find such a downward oligopoly bias.

competitive gains from trade due to reduced markups.

Related literature. Our paper builds on the literature deriving theory-consistent gravity equations. Anderson (1979), Eaton and Kortum (2002), Anderson and van Wincoop (2003), Chaney (2008), Melitz and Ottaviano (2008), Arkolakis, Costinot, and Rodriguez-Clare (2012), Arkolakis, Costinot, Donaldson, and Rodriguez-Clare (2019) and Allen, Arkolakis, and Takahashi (2020) show how to obtain aggregate/industry-level gravity equations from a variety of theoretical frameworks. We contribute to this literature by deriving theory-consistent gravity equations at the firm- and industry level under oligopoly.

Another strand of the literature, surveyed by Head and Mayer (2014), is concerned with the estimation of gravity equations. We contribute to this literature by proposing methods to estimate gravity equations when firms have market power. Anderson and van Wincoop (2003) highlight the importance of controlling for 'multilateral resistance' (i.e., the price index in the destination). Harrigan (1996) was the first to do so using destination fixed effects, an approach that has been followed in most subsequent studies, including the present paper. Santos Silva and Tenreyro (2006) advocate the use of the Poisson Pseudo Maximum Likelihood (PPML) estimator to address a potential bias arising from heteroscedasticity in log-linearized models—an approach that we also follow. An important problem for the estimation of gravity estimations arises from firms self-selecting into export markets. At the firm level, Bas, Mayer, and Thoenig (2017) propose to focus on top exporters that are present in most destinations.[4] At the industry/aggregate level, Helpman, Melitz, and Rubinstein (2008) propose a two-step estimation procedure, which in addition to the standard Heckman correction also controls for the extensive margin of exports. We adopt both approaches in this paper.

Our paper is among the first to use gravity estimates to evaluate the welfare effects of trade policies under oligopoly. Arkolakis, Costinot, and Rodriguez-Clare (2012) identify a class of models with monopolistic or perfect competition in which the trade elasticity is constant and constitutes (in conjunction with countries' trade shares) a sufficient statistic for the welfare gains from trade.[5] Arkolakis, Costinot, Donaldson, and Rodriguez-Clare (2019) extend this approach to a class of monopolistic competition models with variable markups, assuming that productivities are Pareto-distributed. They show empirically that gains from trade are lower when markups are variable rather than constant. While our model does not admit a sufficient statistic, we ask a related but different question: are the gains from trade higher or lower

[4] By focusing on the largest exporters, however, this approach is likely to exacerbate the oligopoly bias, thus making it even more important to control for market power.

[5] See, however, Melitz and Redding (2015) who emphasize the role of micro structure for the gains from trade in models that do not fit the assumptions of Arkolakis, Costinot, and Rodriguez-Clare (2012).

under oligopoly, which features variable markups, compared to monopolistic competition with constant markups? We find significantly larger welfare gains under oligopoly. This is in line with Edmond, Midrigan, and Xu (2015) who calibrate a two-country version of the oligopoly model of Atkeson and Burstein (2008) to assess the gains from trade.[6]

In the last decade, there has been a revived interest in integrating oligopoly into models of international trade, partially building on earlier contributions by the strategic trade policy literature (see Brander, 1995). The by-now dominant framework, which we also adopt, was proposed by Neary (2003) and further developed by Atkeson and Burstein (2008). It features a continuum of oligopolistic industries, implying that firms have market power in their own industry but not in the aggregate. Quantitative papers that build on this framework include Edmond, Midrigan, and Xu (2015) and Gaubert and Itskhoki (2021).[7]

The rest of the paper is organized as follows. In Section 2, we present our theoretical framework and derive oligopoly correction terms for firm- and industry-level gravity equations. In Section 3, we describe the data sources, discuss estimation challenges, and present the empirical results from our firm-level gravity estimations. In Section 4, we repeat these steps for our industry-level gravity estimations. In Section 5, we provide Monte Carlo simulations to evaluate the performance of our estimation procedures. In Section 6, we use a calibrated version of our model to study the welfare gains from trade-cost reductions. Finally, we conclude in Section 7.

2 Gravity Equations under Oligopoly

In this section, we first present the oligopoly model underlying our approach to gravity with granular firms. Next, we derive gravity equations at both the firm and industry level.

2.1 Theoretical Framework

We consider a multi-country world with a continuum of industries, indexed by $z \in \mathcal{Z}$. We denote by $\mathcal{J}_n(z)$ the set of industry-z products sold in country n. We assume that demand

[6] Heid and Stähler (2024) propose an extension of Arkolakis, Costinot, and Rodriguez-Clare (2012)'s formula to evaluate the gains from trade under oligopoly. To recover the necessary parameters, they derive a firm-level gravity equation in oligopoly. However, they estimate it from aggregate trade data, assuming the economy consists of a large number of identical industries, each of which hosts only one firm per country. They also find that the welfare gains from trade liberalization are substantially larger under oligopoly.

[7] Other papers introducing oligopoly into international trade models include Eckel and Neary (2010), Parenti (2018), Breinlich, Nocke, and Schutz (2020), and Head and Mayer (2023).

and inverse demand for any such product $i \in \mathcal{J}_n(z)$ take a CES form:[8]

$$q_{in}(z) = a_{in}(z) p_{in}(z)^{-\sigma(z)} P_n(z)^{\sigma(z)-1} E_n(z) \tag{1}$$
$$p_{in}(z) = a_{in}(z)^{\frac{1}{\sigma(z)}} q_{in}(z)^{-\frac{1}{\sigma(z)}} Q_n(z)^{-\frac{\sigma(z)-1}{\sigma(z)}} E_n(z),$$

where $E_n(z)$ denotes the total expenditure on industry-z products in country n, $a_{in}(z)$ the perceived quality of product i in country n, and $\sigma(z) > 1$ the elasticity of substitution in industry z. The industry-z CES price index and composite commodity in country n are denoted by $P_n(z)$ and $Q_n(z)$, respectively, and given by

$$P_n(z) \equiv \left[\sum_{j \in \mathcal{J}_n(z)} a_{jn}(z) p_{jn}(z)^{1-\sigma(z)} \right]^{\frac{1}{1-\sigma(z)}} \text{ and } Q_n(z) \equiv \left[\sum_{j \in \mathcal{J}_n(z)} a_{jn}(z)^{\frac{1}{\sigma(z)}} q_{jn}(z)^{\frac{\sigma(z)-1}{\sigma(z)}} \right]^{\frac{\sigma(z)}{\sigma(z)-1}}.$$

From now on, we focus on a single industry and drop the index z.[9]

Each product $i \in \mathcal{J}_n$ is offered by a unique firm and produced at constant marginal cost c_{in}. For the firm to sell product i in country n, it incurs iceberg-type trade cost τ_{in} so that its profit from selling q_{in} units in destination n is $\pi_{in} = (p_{in} - \tau_{in} c_{in}) q_{in}$.

Firms compete in quantities in each market n, being able to segment markets perfectly.[10] Under oligopoly, firms take into account the impact of their quantity choices on the CES-composite, Q_n. For what follows, it is useful to generalize further the degree of strategic interaction between firms by introducing a conduct parameter, λ (see Bresnahan, 1989): when firm i increases its output q_{in} by an infinitesimal amount, it perceives the induced effect on Q_n to be equal to $\lambda \partial Q_n / \partial q_{in}$. Under monopolistic competition, the conduct parameter λ takes the value of zero, whereas it is equal to one under Cournot competition. The first-order condition of profit maximization of firm i in destination n is given by

$$0 = \frac{\partial \pi_{in}}{\partial q_{in}} = \frac{E_n}{Q_n^{\frac{\sigma-1}{\sigma}}} a_{in}^{\frac{1}{\sigma}} \frac{\sigma-1}{\sigma} q_{in}^{-\frac{1}{\sigma}} - \frac{\sigma-1}{\sigma} \lambda \frac{\partial Q_n}{\partial q_{in}} \frac{E_n a_{in}^{\frac{1}{\sigma}} q_{in}^{\frac{\sigma-1}{\sigma}}}{Q_n^{\frac{\sigma-1}{\sigma}+1}} - \tau_{in} c_{in}$$

[8] As is well known, such a demand function could be derived, for example, from a two-tier utility function, with Cobb-Douglas at the upper tier and CES at the lower tier. In Section 5, we use a quasi-linear version of that utility function to generate the same demand.

[9] The framework we lay out here can be viewed as being general equilibrium. However, as we focus on a given equilibrium and do not conduct comparative statics at the aggregate level, we refrain from explicitly closing the model by writing down factor market-clearing conditions and endogenizing consumer income. Closing the model would be straightforward. For example, we could assume that the demand system has been derived from the maximization of a two-tier utility function, with Cobb-Douglas at the upper tier, and then assign a labor endowment to each country, assume that all costs are incurred in terms of origin-country labor, choose labor in a reference country as the numeraire, and assume that profits and tariff revenues are distributed lump sum to domestic consumers.

[10] We focus on quantity competition here and present results for price competition in Online Appendix C.

$$= \frac{\sigma - 1}{\sigma} p_{in} \left(1 - \lambda s_{in}\right) - \tau_{in} c_{in}, \tag{2}$$

where

$$s_{in} \equiv \frac{a_{in}^{\frac{1}{\sigma}} q_{in}^{\frac{\sigma-1}{\sigma}}}{\sum_{j \in \mathcal{J}_n} a_{jn}^{\frac{1}{\sigma}} q_{jn}^{\frac{\sigma-1}{\sigma}}} \tag{3}$$

is the market share of firm i in destination n.

Rearranging terms in equation (2) yields firm i's optimal markup in destination n:

$$\mu_{in} = \frac{1}{\sigma} + \lambda \frac{\sigma - 1}{\sigma} s_{in}, \tag{4}$$

where $\mu_{in} \equiv \left(p_{in} - \tau_{in} c_{in}\right) / p_{in}$ is the Lerner index. Under monopolistic-competition conduct ($\lambda = 0$), the usual constant markup $1/\sigma$ obtains. If instead $\lambda > 0$, then markups are no longer constant and depend positively on market shares. We will make use of the additional flexibility afforded by the conduct parameter λ in Section 2.3, but for now, we assume Cournot conduct and set $\lambda = 1$.

2.2 Firm-Level Gravity in Oligopoly

From the definition of the Lerner index, firm i's price in market n is $p_{in} = c_{in}\tau_{in}/(1 - \mu_{in})$. Using equation (1), the value of its sales can be written as

$$r_{in} = p_{in} q_{in} = \left(\frac{c_{in}\tau_{in}}{1 - \mu_{in}}\right)^{1-\sigma} a_{in} P_n^{\sigma-1} E_n. \tag{5}$$

We log-linearly decompose the quality and cost terms as $\log a_{in} = \varepsilon_i^a + \varepsilon_n^a + \varepsilon_{in}^a$ and $\log c_{in} = \varepsilon_i^c + \varepsilon_n^c + \varepsilon_{in}^c$, respectively. We further decompose trade costs as $\log \tau_{in} = \beta X_{in} + \varepsilon_i^\tau + \varepsilon_n^\tau + \varepsilon_{in}^\tau$ where X_{in} includes variables with bilateral variation such as (log) distance, common language, or dummies for the presence of trade agreements or currency unions.

Taking the logarithm of equation (5) yields the firm-level gravity equation

$$\log r_{in} = \xi_n + \zeta_i + \beta(1 - \sigma) X_{in} + (\sigma - 1) \log\left(1 - \mu_{in}\right) + \varepsilon_{in}, \tag{6}$$

where ξ_n and ζ_i summarize destination- and firm-specific terms, and $\varepsilon_{in} = \varepsilon_{in}^a + (1 - \sigma)\left(\varepsilon_{in}^c + \varepsilon_{in}^\tau\right)$ collects the dyadic unobservables. Obtaining a consistent estimate of $\beta(1-\sigma)$, the coefficients on the bilateral variables, is a key objective of much of the gravity literature.

If the data were generated from monopolistic competition, the markup term involving μ_{in} in equation (6) would be constant and could be subsumed in ζ_i. In that case, estimation of equation (6) would yield a consistent estimate of $\beta(1-\sigma)$, provided that firm and destination fixed effects (ξ_n and ζ_i) are included and that the usual identifying assumption made in the gravity literature hold.[11] To back out β requires, of course, knowledge of σ—either from another source or from the gravity estimation itself, provided X_{in} contains bilateral tariffs.[12]

Suppose instead that the data are generated from oligopoly, which is likely to be the case in many industries. Then, the markup term—which is a function of firms' market shares—will go into the error term, introducing a correlation between the latter and the regressors of interest, X_{in}. For example, to the extent that firms face larger variable trade costs in more-distant markets, their market shares are lower there, ceteris paribus. (This is indeed borne out by the data; see Table 2 below.) Hence, firms charge lower markups in such destinations, implying a positive correlation between distance and the omitted variable, $\log(1-\mu_{in})$, and thus a bias in the estimate of the distance coefficient.[13,14]

The solution that we now propose allows us to obtain consistent estimates of both the structural parameter β (or $\beta(1-\sigma)$) and the effect of changes in gravity variables on trade flows. Specifically, suppose that we have data on market shares and also an estimate of σ. Computing firm i's markup in destination n as

$$\widehat{\mu}_{in} = \frac{1}{\widehat{\sigma}} + \frac{\widehat{\sigma}-1}{\widehat{\sigma}} s_{in},$$

we can then "purge" the observed trade flows from oligopolistic market power effects as

[11] For least-squares estimation of the log-linearized gravity equation, the identifying assumption is $\mathbb{E}\left[\varepsilon_{in}|X_{in},\xi_n,\zeta_i\right]=0$. This assumption does not rule out correlations between the bilateral variables and taste, production and trade cost shocks working through the firm- and destination-level components (ε_i^a, ε_n^a, ε_i^c, ε_n^c, ε_i^τ and ε_n^τ). Such correlations are not a problem, as these components can be controlled for through firm and destination fixed effects. If the data contain a time dimension, one can also allow for time-invariant bilateral elements in the error term which can be captured through bilateral fixed effects—as is standard, e.g., in the literature on the trade effects of preferential trade agreements (see Baier and Bergstrand, 2007).

[12] In the latter case, the coefficient on $\log(1+t_{in})$, where t_{in} is the ad valorem tariff, can be shown to be equal to $1-\sigma$.

[13] As markups vary at the firm-destination level, their variation cannot be controlled for by firm and destination fixed effects. The inclusion of firm-destination fixed effects would make it impossible to identify separately the effect of key regressors of interest such as distance, tariffs or dummy variables for trade agreements. Having a time dimension in the data would not help either because markups would then vary by firm, destination, and time.

[14] If firms compete in prices instead of quantities, then the equilibrium markup of firm i in destination n satisfies $\mu_{in}=1/(\sigma-(\sigma-1)s_{in})$ (see Online Appendix C.1 for a theoretical treatment of the price competition case). A first-order approximation around $s_{in}\simeq 0$ yields

$$\mu_{in} \simeq \frac{1}{\sigma} + \frac{\sigma-1}{\sigma^2} s_{in},$$

implying that markups are less sensitive to market shares under price competition than under Cournot competition (compare equation (4)). We therefore expect the oligopoly bias to be smaller under price competition. Our empirical results confirm this intuition (see Table 3 below and Table A in Online Appendix C.2).

follows:

$$\log \widetilde{r}_{in} \equiv \log r_{in} - (\widehat{\sigma} - 1) \log (1 - \widehat{\mu}_{in}). \tag{7}$$

Combining equations (6) and (7), we obtain a standard gravity equation for the purged trade flows:

$$\log \widetilde{r}_{in} = \xi_n + \zeta_i + \beta(1-\sigma) X_{in} + \varepsilon_{in}. \tag{8}$$

Under the usual identifying assumption (see footnote 11), the empirical specification in equation (8) allows us to obtain a consistent estimate of $\beta(1-\sigma)$ and thus of β.[15,16]

Above, we assumed access to a prior estimate of σ. Such a prior estimate is not required if X_{in} contains information on tariffs (t_{in}). In Appendix B, we show that the standard identifying assumption, $\mathbb{E}[\varepsilon_{in}|X_{in}, t_{in}, \xi_n, \zeta_i] = 0$, can then be used to form moment conditions and estimate σ and β by GMM.

In addition to the structural parameters, trade economists are often interested in estimating the effect of changes in the gravity variables on trade flows, holding fixed (monadic) firm and destination characteristics; that is, in $\mathbb{E}[\nabla_{X_{in}} \log r_{in}|X_{in}, \xi_n, \zeta_i]$. In Appendix A.3, we show that the partial effect of the gravity variables, conditional on s_{in}, is given by:[17]

$$\mathbb{E}[\nabla_{X_{in}} \log r_{in}|X_{in}, s_{in}, \xi_n, \zeta_i] = \beta(1-\sigma) \frac{1}{1 + (\sigma - 1)\frac{s_{in}}{1-s_{in}}}. \tag{9}$$

This partial effect—which incorporates the firm's endogenous markup adjustment—can be computed using our estimates of β and σ, and data on market shares. The average partial effect can then be computed by integrating over the distribution of market shares. That is, our solution to the oligopoly bias allows us to obtain both the structural parameter and the effect of gravity variables on trade flows. By contrast, regressing observed trade flows on the bilateral variables and fixed effects does not, in general, yield a consistent estimate of the average partial effect of gravity variables on trade flows.[18]

[15] Note the parallel to the literature on trade and quality which uses a similar approach to correct export values or quantities (e.g., Khandelwal, Schott, and Wei, 2013).

[16] The alternative approach of writing the markup term explicitly as a function of market shares and including it as a regressor does not deliver a consistent estimate of the structural parameters. To see this, note that equation (6) can be rewritten as

$$\log r_{in} = \xi_n + \zeta_i + \beta(1-\sigma) X_{in} + (\sigma - 1) \log (1 - s_{in}) + \varepsilon_{in},$$

where we have used equation (4). As $\log (1 - s_{in})$ is correlated with both X_{in} and ε_{in}, least-square estimation does not yield a consistent estimate of $\beta(1-\sigma)$.

[17] As is standard in the gravity literature, the price index in the destination country is held fixed when computing the effect.

[18] Due to markup adjustments, the conditional expectation $\mathbb{E}[\log r_{in}|X_{in}, \xi_n, \zeta_i]$ is highly nonlinear in (X_{in}, ξ_n, ζ_i). Hence, the estimate from such a regression would be a weighted average of the slopes

2.3 Industry-Level Gravity in Oligopoly

We now turn to gravity at the industry level. We first analyze the equilibrium in a given market using an aggregative games approach (Nocke and Schutz, 2018; Anderson, Erkal, and Piccinin, 2020). We then leverage Nocke and Schutz (2024)'s approximation techniques to derive an industry-level gravity equation that accounts for oligopolistic behavior.

An aggregative games approach to industry equilibrium. Consider industry z in destination n. Dropping reference to z to ease notation, we define the market-level aggregator H_n as

$$H_n \equiv Q_n^{\frac{\sigma-1}{\sigma}} = \sum_{j\in\mathcal{J}_n} a_{jn}^{\frac{1}{\sigma}} q_{jn}^{\frac{\sigma-1}{\sigma}}$$

and firm i's type T_{in}, a measure of quality-adjusted productivity, as

$$T_{in} \equiv a_{in}^{\frac{1}{\sigma}} \left(\frac{E_n}{c_{in}\tau_{in}} \frac{\sigma-1}{\sigma} \right)^{\frac{\sigma-1}{\sigma}}. \tag{10}$$

Combining these definitions with equations (2) and (3), we obtain:

$$1 - \lambda s_{in} = s_{in}^{\frac{1}{\sigma-1}} \left(\frac{H}{T_{in}} \right)^{\frac{\sigma}{\sigma-1}}. \tag{11}$$

As the left-hand side is non-increasing in s_{in} and the right-hand side is strictly increasing in s_{in}, the equation has a unique solution in s_{in}, denoted $S(T_{in}/H_n, \lambda)$—the *market-share fitting-in function*. It is easily verified that $S(\cdot,\cdot)$ is strictly increasing in its first argument and strictly decreasing in its second.

The equilibrium level of the aggregator, $H^*(\lambda)$, is pinned down by market shares adding up to one:

$$\sum_{i\in\mathcal{J}_n} S\left(\frac{T_{in}}{H_n}, \lambda\right) = 1. \tag{12}$$

The uniqueness of the solution follows by the strict monotonicity of the market-share fitting-in function.

To summarize:

Proposition 1. *In each destination market n, and for any conduct parameter λ, there exists a unique equilibrium in quantities. The equilibrium aggregator level $H_n^*(\lambda)$ is the unique solution to equation* (12)*. Each firm i's equilibrium market share is $s_{in}^*(\lambda) = S(T_{in}/H_n^*(\lambda), \lambda)$, where $S(T_{in}/H_n^*(\lambda), \lambda)$ is the unique solution to equation* (11)*.*

$\mathbb{E}[\nabla_{X_{in}} \log r_{in} | X_{in}, \xi_n, \zeta_i]$, but with weights being generally different from the density of (X_{in}, ξ_n, ζ_i) (see, e.g., Yitzhaki, 1996; Angrist and Pischke, 2009).

Proof. See Appendix A.1 □

The first-order approach to industry-level gravity. Let $\mathcal{J}_{on} \subsetneq \mathcal{J}_n$ denote the subset of exporters from country o that sell in the destination market n. Their aggregate exports to market n are given by $s^*_{on}E_n$, where $s^*_{on}(\lambda) \equiv \sum_{i\in\mathcal{J}_{on}} s^*_{in}(\lambda)$. We are interested in these aggregate exports when firms compete in a Cournot fashion, i.e., when $\lambda = 1$. Unfortunately, there is no closed-form solution to $s^*_{on}(1)$. Our approach therefore entails approximating it.

As we show in the following, the approximation relies on two Herfindahl indices, namely the HHI of all firms selling in the destination market n,

$$\mathrm{HHI}_n(\lambda) \equiv \sum_{j\in\mathcal{J}_n} \left(s^*_{jn}(\lambda)\right)^2,$$

and the HHI among the exporters in country o that sell in the destination market n,

$$\mathrm{HHI}_{on}(\lambda) \equiv \sum_{j\in\mathcal{J}_{on}} \left(\frac{s^*_{jn}(\lambda)}{s^*_{on}(\lambda)}\right)^2.$$

We obtain:

Proposition 2. *At the first order, in the neighborhood of $\lambda = 0$ (monopolistic-competition conduct), the logged joint market share in destination n of the firms from origin o is given by*

$$\log s^*_{on}(\lambda) = \log s^*_{on}(0) + (\sigma - 1)\Big[\mathrm{HHI}_n(\lambda) - s^*_{on}(\lambda)\,\mathrm{HHI}_{on}(\lambda)\Big]\lambda + o(\lambda).$$

Proof. See Appendix A.2. □

The proposition shows that the joint market share of the exporters from country o differs from the one that would obtain under monopolistic competition by a market-power term that takes account of both the overall concentration in the destination market and the concentration among the country-o exporters.

This result motivates the following approximation:

$$\log s^*_{on}(0) \simeq \log s^*_{on}(1) - (\sigma - 1)\Big[\mathrm{HHI}_n(1) - s^*_{on}(1)\,\mathrm{HHI}_{on}(1)\Big].$$

Thus, the export flow that would obtain under monopolistic competition is approximately given by

$$\underbrace{\log(E_n) + \log s^*_{on}(1)}_{\log r_{on}} - (\sigma - 1)\Big[\mathrm{HHI}_n(1) - s^*_{on}(1)\,\mathrm{HHI}_{on}(1)\Big],$$

where r_{on} is the (actually observed) export flow under oligopoly. As the HHI_n term will be subsumed in the destination fixed effect, we define

$$\log \widetilde{r}_{on} \equiv \log r_{on} + (\sigma - 1) s_{on}\, \text{HHI}_{on} \tag{13}$$

as the value of the export flow from o to n purged from market-power effects, which can be computed with data on r_{on}, s_{on} and HHI_{on}, and an estimate of σ.[19]

Next, we derive a gravity equation for oligopoly-corrected trade flows $\log \widetilde{r}_{on}$. To do so, we impose the following structure on the quality, marginal-cost and trade-cost terms:

$$\begin{aligned} \log a_{in} &= \log a_i + \varepsilon_o^a + \varepsilon_n^a + \varepsilon_{on}^a, \\ \log c_{in} &= \log c_i + \varepsilon_o^c + \varepsilon_n^c + \varepsilon_{on}^c, \\ \log \tau_{in} &= \beta X_{on} + \varepsilon_o^\tau + \varepsilon_n^\tau + \varepsilon_{on}^\tau. \end{aligned}$$

Combining this with equation (5) (with $\mu_{in} = 1/\sigma$), adding up over all the exporters from o to n, and taking the logarithm, yields a gravity equation of the following form:

$$\log \widetilde{r}_{on} = \xi_o + \zeta_n + \beta(1 - \sigma) X_{on} + \phi_{on} + \varepsilon_{on}, \tag{14}$$

where ξ_o is an origin fixed effect, ζ_n is a destination fixed effect,

$$\phi_{on} \equiv \log \sum_{j \in \mathcal{J}_{on}} a_j c_j^{1-\sigma},$$

and

$$\varepsilon_{on} \equiv \varepsilon_{on}^a + (1 - \sigma)(\varepsilon_{on}^c + \varepsilon_{on}^\tau).$$

If the set of exporters from origin o were the same in all destinations n (i.e., if $\mathcal{J}_{on}$ were independent of n), then the term ϕ_{on} would be subsumed into the origin fixed effect. In that case, regressing $\log \widetilde{r}_{on}$ on origin and destination fixed effects, and the bilateral variables X_{on} would yield a consistent estimate of $\beta(1-\sigma)$, provided the usual identifying assumption $\mathbb{E}\left[\varepsilon_{on} | X_{on}, \xi_o, \zeta_n\right] = 0$ holds.

If, instead, the set $\mathcal{E}_{on}$ does depend on n because of self-selection into export destinations, then ϕ_{on} is no longer absorbed by the origin fixed effect, and is likely to be correlated with

[19] If firms compete in prices instead of quantities, then the market power term in equation (13) becomes $[(\sigma-1)/\sigma] s_{on}\, \text{HHI}_{on}$ (see Proposition B in Online Appendix C.1). That is, the oligopoly correction term under price competition is equal to that under quantity competition divided by σ. We therefore expect the oligopoly bias to be smaller under price competition, as was the case for firm-level regressions (see footnote 14). Our empirical results confirm this intuition (see Table 8 below and Table B in Online Appendix C.2).

X_{on}. In Section 4 below, we discuss how to address such self-selection issues and obtain a consistent estimate of $\beta(1-\sigma)$.

We can also approximate the effect of changes in gravity variables on trade flows *for a given set of exporters from o to n*. We focus on this *intensive-margin* effect for two reasons. First, this is also the effect of interest at the firm level. Second, the *overall* effect of trade costs on trade flows, which would include the extensive-margin effect through self-selection into exporting, is generally not constant, as competition is oligopolistic and productivities are not (necessarily) Pareto-distributed. In Appendix A.3, we derive the following approximation (around small market shares) of the intensive-margin effect:

$$\mathbb{E}[\nabla_{X_{on}} \log r_{on} | X_{on}, s_{on}, \mathrm{HHI}_{on}, \xi_o, \zeta_n] \simeq \beta(1-\sigma)\left[1-(\sigma-1)s_{on}\,\mathrm{HHI}_{on}\right]. \tag{15}$$

This expression can be computed using the estimate of $\beta(1-\sigma)$ and data on aggregate market shares and Herfindahl indices.

3 Empirical Implementation: Firm-Level Gravity

In this section, we show how to empirically implement our gravity-estimation approach at the firm level. Our empirical specification is

$$\log \widetilde{r}_{inzt} = \xi_{nzt} + \zeta_{izt} + \beta(1-\sigma)X_{in} + \varepsilon_{inzt}, \tag{16}$$

where

$$\log \widetilde{r}_{inzt} \equiv \log r_{inzt} - (\widehat{\sigma}-1)\log\left(1-s_{inzt}\right), \tag{17}$$

corresponding to equations (8) and (7) in Section 2.2 above, except that we have made here the industry (z) and time (t) dimensions explicit and made use of equation (4). Due to data limitations explained below, we focus on distance as our only gravity variable, so that X_{in} boils down to the scalar $\log(\mathrm{dist}_{on})$, where dist_{on} is the distance between firm i's origin o and destination n. In the following, we discuss estimation challenges, present our data, run gravity regressions with and without oligopoly correction, and investigate under what circumstances ignoring oligopolistic behavior leads to quantitatively important biases.

Estimation Challenges. A first issue is how to control for destination fixed effects ξ_{nzt} in a setting with firm-level export data. With export data from a single origin country, we would not be able to separate the impact of bilateral variables from the destination fixed effects.[20] To address this issue, we follow Bas, Mayer, and Thoenig (2017) by combining two datasets on the exports of French and Chinese firms, respectively.

Secondly, we have to address self-selection issues, as most firms export only to a subset of possible destinations. When estimating equation (16), observations with zero trade flows drop out. In the presence of export fixed cost, there is selection into exporting in our model: firms selling in more distant foreign markets will be more likely to have received a favorable taste, productivity, or trade-cost shock for that destination, allowing them to operate in this more difficult environment. As a consequence, the conditional expectation $\mathbb{E}\left[\varepsilon_{inzt}|X_{in}, r_{inzt} > 0\right]$ is likely to depend on X_{in}. To address this issue, we adapt an approach proposed by Bas, Mayer, and Thoenig (2017) and restrict our estimation sample to the largest three French firms and the largest three Chinese firms in each industry, as measured by total industry-level exports, added up over all destinations. As those firms are generally very productive, produce high-quality products (low ε^c_{izt} and/or high ε^a_{izt}), or use low-cost market-access technologies (low ε^τ_{izt}), they are likely to serve most destinations, so that the destination-specific shocks (ε^c_{inzt}, ε^a_{inzt}, and ε^τ_{inzt}) do not play an important role in their market entry decisions. We acknowledge that this is an imperfect solution but our simulation evidence presented in Section 5 shows that focusing on top exporters does indeed substantially reduce selection bias. Moreover, we show that our results are very similar when using only the top exporter or top-5 exporters from France and China.

Third, as shown by Santos Silva and Tenreyro (2006), in the presence of heteroscedasticity the log-linearized gravity equation yields inconsistent estimates of $\mathbb{E}\left[\widetilde{r}_{inzt}|X_{in}, \xi_{nzt}, \zeta_{izt}\right]$, where

$$\widetilde{r}_{inzt} = \exp[\xi_{nzt} + \zeta_{izt} + \beta(1-\sigma)X_{in}]\exp(\varepsilon_{inzt}). \tag{18}$$

To see this, note that if $\mathrm{Var}(\exp(\varepsilon_{inzt})|X_{in})$ depends on X_{in}, then so does $\mathbb{E}\left[\varepsilon_{inzt}|X_{in}\right]$. A solution to this problem is to estimate the gravity equation (18) by PPML in multiplicative form, which also allows us to include zero trade flows in our estimation sample. Recent computational advances in PPML estimation (e.g., Correia, Guimaraes, and Zylkin, 2019) make it possible to include the large number of fixed effects required in our setting.

Finally, the oligopoly correction term for firm-level gravity (see equation (17)) requires estimates of σ. In the main part of the paper, we consider values of σ equal to 4, 5, and 6—in line with standard estimates in the trade literature.[21]

[20] For example, if we used data on the exports of French firms only, we would not be able to distinguish whether firms' exports to a given destination are high because France and the country in question are close to each other or because of other destination-specific factors such as a high price index or expenditure level.

[21] Estimates of σ are usually in the range from 4 to 6. For example, Bas, Mayer, and Thoenig (2017) find values of σ ranging from 4.2 to 7; Gaubert and Itskhoki (2021) obtain an "imprecisely estimated" $\sigma = 4.9$; Breinlich, Nocke, and Schutz (2020) calibrate $\sigma = 5.2$ in the median 5-digit industry. If our data exhibited

Data. We use annual firm-level export data for French and Chinese exporters provided by the two countries' customs authorities for the years 2000–2010. In each dataset, we observe all the products a firm exports and all the destinations it serves, and the value of the underlying flows. Although both datasets record export data at the 8-digit level, we need to aggregate this information up to the 6-digit level of the Harmonised System (HS), which is the most disaggregate level at which the two national classifications are comparable.

To compute market shares, defined as the ratio of export value to absorption, we combine our firm-level data with absorption data at the HS 6-digit level (or close to it) from Eurostat's PRODCOM database.[22] The downside of using PRODCOM is that absorption data is only available for European countries. As a result, there is insufficient variation to include, in addition to distance, regressors such as dummies for common language or policy-related variables (e.g., membership in a free-trade agreement and bilateral tariffs).[23]

After combining our data sources, we end up with information on export values, export quantities, and market shares for 31 European destinations, 1,864 industries and around 250,000 exporters for the period 2000–2010.[24,25] We source information on bilateral distance from CEPII.[26]

tariff variation by destination, we could use the approach described in Appendix B to estimate σ directly from the gravity equation (see also Head and Mayer, 2023). In the absence of such data, we show in an earlier version of this paper (Breinlich, Fadinger, Nocke, and Schutz, 2023) how to estimate σ by adapting the estimation procedure of Feenstra (1994) and Broda and Weinstein (2006) to firm-level data and oligopolistic competition.

[22]Absorption, defined as domestic production + imports − exports, is the counterpart to $E_{nt}(z)$ in our model. In principle, this information is available at the HS 6-digit level but issues such as classification changes over time often require aggregation to higher levels. The original classification of the PRODCOM data is the 8-digit CN classification, which changes almost every year. We apply the procedure developed by Van Beveren, Bernard, and Vandenbussche (2012) to map the CN classification to an artificial HS classification, "HS 6-digit plus", that is comparable over time and compatible with the 6-digit HS classification. The idea is to aggregate both trade and PRODCOM data as little as possible and as much as required to guarantee a one-to-one mapping between them. See their paper for an in-depth discussion of the procedure.

[23]The destination countries in our sample were either EU member states or had implemented free-trade agreements (FTAs) with the EU before 2000 and therefore had no tariffs on EU imports. By contrast, China did not have any FTAs with countries in our sample and EU external tariffs for imports from China only had variation across industries. Thus, all variation in the FTA dummy or in tariffs would be absorbed by our firm-industry-year fixed effects. Likewise, there is insufficient variation to include an indicator for common language.

[24]The export destinations are: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Norway, Poland, Portugal, Romania, Serbia, Slovakia, Slovenia, Spain, Sweden, Turkey, and the United Kingdom.

[25]Possibly because of measurement issues in PRODCOM, we occasionally observe instances where absorption is smaller than a firm's export value, resulting in market shares larger than one; in such cases, we winsorize market shares to 0.95.

[26]Specifically, we use the population-weighted distance measure `distw` from the CEPII database.

Descriptive Statistics. Firm-level market shares are the key determinant of our oligopoly correction term. In Table 1, column (1) presents summary statistics on the firm-level market shares for the French and Chinese exporters. The average market share across the approximately 14 million firm-destination-industry-year combinations in our data is small (0.4%) and the median is even smaller (around 0.01%). At the 95th percentile, the firm-level market share is 1.12%. Clearly, the typical firm in our data does not enjoy much market power.

However, this does not imply that correcting firm-level exports for oligopoly forces will not matter quantitatively, as estimation results could be substantially biased by a small number of exporters with large market shares. Columns (2)–(4) focus on such firms. Column (2) shows descriptive statistics for the top exporters (i.e., for any given 6-digit industry and year, the French firm and the Chinese firm with the largest total export value). The average top-exporter market share is around 6%, substantially larger than the average exporter's market share. Moreover, at the 95th percentile the top firm enjoys a market share of almost 30%. Columns (3) and (4) present summary statistics on the market shares and cumulative market shares of the top-3 French and the top-3 Chinese exporters. The main takeaway is that, in a significant minority of destination markets, the largest French and Chinese exporters command substantial market shares.

Table 1: Summary Statistics for the Market Shares of French and Chinese Exporters

	(1) All Exporters	(2) Top Exporters	(3) Top 3 Exporters	(4) Top 3 Exporters (Cumulative)
Mean	0.40%	6.00%	3.88%	7.30%
5th pctile	0.00007%	0.01%	0.006%	0.03%
10th pctile	0.0004	0.03	0.02%	0.09%
Median	0.01%	1.21%	0.65%	2.05%
90th pctile	0.44%	15.72%	9.20%	19.36%
95th pctile	1.12%	28.96%	18.04%	33.44%
Observations	14,009,005	276,718	708,409	708,409

Notes: Table shows summary statistics for strictly positive market shares of French and Chinese exporters for the years 2000–2010. The unit of observation is at the firm-destination-industry-year level.

Table 2 provides the results from regressing the logged market shares of the top-3 exporters on logged distance, controlling for firm-industry-year and destination-industry-year fixed effects. As expected, the coefficient on log-distance is strongly negative and statistically significant, which is consistent with an oligopoly bias in standard gravity estimation.

Gravity Estimation Results. All regressions are run on the sample of the top-3 French and top-3 Chinese exporters for each 6-digit HS industry. As a first step, we pool all industries

Table 2: Regressing Log Market Shares on Log Distance

	OLS
log dist.	-0.239***
	(0.013)
Observations	708,392
R^2	0.02
Firm-ind.-year FE	YES
Ind.-dest-year FE	YES

Note: Firm-level data. Results for top 3 exporters. Standard errors in parentheses, clustered at the destination-year level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

and years and estimate equation (18) by PPML. We do so both on actual export flows and on oligopoly-corrected export flows, as defined in equation (17).

Table 3: Firm-Level Gravity Estimates

	(1)	(2)	(3)	(4)
Method	PPML w/o corr	PPML w/ corr $\sigma = 4$	PPML w/corr $\sigma = 5$	PPML w/ corr $\sigma = 6$
log dist.	-0.874***	-1.492***	-1.518***	-1.528***
	(0.021)	(0.210)	(0.220)	(0.223)
Obs.	11,955,786	11,955,786	11,955,786	11,955,786
(Pseudo) R^2	0.14	0.27	0.28	0.29
Firm-ind.-year FE	YES	YES	YES	YES
Ind.-dest.-year FE	YES	YES	YES	YES

Notes: Firm-level data, pooled across industries and years. Results for top-3 exporters, without and with oligopoly correction. Standard errors in parentheses, clustered at the destination-year level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3 reports the estimated coefficient on log distance, $\beta(1-\sigma)$. Column (1) presents the PPML estimate for the specification without oligopoly correction. The value of -0.874 almost exactly corresponds to the median estimate of -0.89 in Head and Mayer (2014)'s meta study of 2508 gravity estimates from 159 papers (see their Table 3.4).[27] Columns (2)–(4) show the results including the oligopoly correction, for different values of σ. Regardless of the value of σ, the absolute value of the estimated coefficient on log distance is more than 70% larger than without oligopoly correction, ranging from 1.492 (for $\sigma = 4$) to 1.528 (for $\sigma = 6$).

[27] By contrast, estimating the gravity equation without oligopoly correction by OLS yields an implausibly low estimate of -0.232, indicating a strong heteroscedasticity bias; see Table C in Online Appendix D. Estimating the same equation without oligopoly correction by PPML, but dropping the observations with zero trade flows, also yields an implausibly low estimate of -0.410; see Table D in Online Appendix D. Our Monte Carlo simulations in Section 5 confirm the presence of such attenuation biases when using OLS and PPML without zeroes; see footnote 37.

The implied distance elasticity of trade costs, β, varies from 0.175 (for $\sigma = 6$) to 0.291 (for $\sigma = 4$) without oligopoly correction, and from 0.306 (for $\sigma = 6$) to 0.497 (for $\sigma = 4$) with correction.[28]

These empirical results confirm our theoretical insight that the trade elasticity with respect to distance suffers from a substantial attenuation bias in the order of 42%. This bias arises because firms systematically reduce their markups in markets where they face higher variable trade costs and thus have lower market shares. As a consequence, export values decrease by less than they would have decreased under constant markups.

We now turn to estimating gravity equations separately for each of the 78 HS 2-digit sectors, pooling observations across 6-digit industries within a given 2-digit sector. Table 4 reports the median of the estimated distance coefficient, both with and without oligopoly correction, and summary statistics on the distribution of the resulting oligopoly bias across sectors. Note that, without oligopoly correction, the median estimated distance coefficient is positive and, thus, has the wrong sign.[29] By contrast, regardless of σ, the median estimated coefficient has the correct (negative) sign and is of similar magnitude to the estimate from the pooled regression. For example, for $\sigma = 5$, the median estimated coefficient is -1.347, compared to -1.518 when pooling observations from all sectors. For this specification, the median value of the absolute percentage bias is 96%; in 10% of sectors, the bias is 160% or larger.[30] Thus, in many industries, the oligopoly bias is much larger than suggested by the estimates from the pooled regressions.

Table 4: Firm-level Gravity Estimates by 2-digit Sector

Median est coefficient	w/o corr	w/ corr $\sigma = 4$	w/ corr $\sigma = 5$	w/corr $\sigma = 6$
log dist.	0.508	-0.840	-1.347	-1.225
abs. pct. bias (10th pctile)		18%	10%	17%
abs. pct. bias (median)		94%	96%	97%
abs. pct. bias (90th pctile)		234%	160%	134%

Notes: Firm-level data. Table shows summary statistics on the distribution of estimated coefficients by 2-digit HS sector for top-3 exporters.

[28] As shown in Table E in Online Appendix D, very similar results obtain when using (i) only the top French exporter and the top Chinese exporter and (ii) the top-5 French exporters and the top-5 Chinese exporters.

[29] Surprisingly, firm-level gravity regressions are rarely run sector by sector. The only exception we are aware of is Bas, Mayer, and Thoenig (2017), who do not report the coefficient on distance. Thus, we are unable to compare our estimates with analogous distance elasticities from the literature.

[30] The absolute percentage bias is defined as the absolute value of $(\widehat{\beta}_{\text{w/o corr}} - \widehat{\beta}_{\text{w/ corr}})/\widehat{\beta}_{\text{w/ corr}}$.

4 Empirical Implementation: Industry-Level Gravity

We now show how to empirically implement our gravity-estimation approach at the industry level. Our empirical specification is

$$\log \widetilde{r}_{onz} = \xi_{oz} + \zeta_{nz} + \beta(1-\sigma)\log(\text{dist}_{on}) + \eta_{onz}, \tag{19}$$

where

$$\log \widetilde{r}_{onz} = \log r_{onz} + (\widehat{\sigma}-1)s_{onz}\,\text{HHI}_{onz}\,. \tag{20}$$

This corresponds to equations (14) and (13) above, with the industry (z) dimension being made explicit, $\log(\text{dist}_{on})$ being our only gravity variable (for the reasons explained in Section 3), and $\eta_{onz} \equiv \phi_{onz} + \varepsilon_{onz}$. We refrain from introducing a time index t because, in the estimations below, we confine attention to data from the year 2010 for computational reasons.

As in the previous section, we now turn to a discussion of the estimation challenges. We then briefly describe our data, run gravity regressions with and without oligopoly correction, and investigate under what circumstances ignoring oligopolistic behavior leads to quantitatively important biases.

Estimation Challenges. The self-selection of firms into export markets again poses problems for a consistent estimation of the structural parameter $\beta(1-\sigma)$. If in each origin country o there were a single firm choosing whether to enter any given destination country n, we would have the same sample selection problem as at the firm level. While the ϕ_{onz}-term would be subsumed into the origin fixed effect, the conditional expectation $\mathbb{E}\left[\varepsilon_{onz}|\widetilde{r}_{onz}>0,\xi_{oz},\zeta_{nz},\text{dist}_{on}\right]$ would depend on dist_{on}: the observation that a firm is exporting to a remote market is likely to be the result of that firm having received a favorable ε_{onz}-shock. With multiple potential exporters, this problem remains. In addition to sample selection, however, a potential extensive-margin bias arises whenever the set of exporters from o varies with n, so that the ϕ_{onz}-term can no longer be subsumed into the origin fixed effect. In particular, the conditional expectation $\mathbb{E}\left[\phi_{onz}|\widetilde{r}_{onz}>0,\xi_{oz},\zeta_{nz},\text{dist}_{on}\right]$ is likely to depend on dist_{on}, as a larger number of firms would presumably find it profitable to export to nearby destinations. Summarizing, while the sample-selection bias tends to lead to an underestimation of the effect of distance on trade, the extensive-margin bias tends to result in an overestimation.

To alleviate both biases, we apply the two-step procedure developed by Helpman, Melitz, and Rubinstein (2008) (henceforth, HMR) to the oligopoly-corrected trade flows. The first step consists in estimating a Probit model of whether positive trade flows between o and n

are observed. The regressors are origin and destination fixed effects, $\log \text{dist}_{on}$, and bilateral variables that are likely to affect the fixed export cost but not variable trade costs. For the latter variables, we follow HMR and use: (i) A dummy equal to one if business startup time is above the median in both o and n; and (ii) a dummy equal to one if business startup cost is above the median in both o and n. The estimated conditional probability of observing positive trade flows is denoted $\widehat{\rho}_{onz}$.

The second step consists in estimating the following gravity equation:

$$\log \widetilde{r}_{onz} = \zeta_{oz} + \xi_{nz} + \beta(1-\sigma)X_{onz} + P(\log \hat{Z}_{onz}) + \omega\widehat{\lambda}_{onz} + \eta_{onz}, \tag{21}$$

where $\widehat{\lambda}_{onz}$ is the inverse Mills ratio from the first step, $\log \hat{Z}_{onz}$ is the $\widehat{\rho}_{onz}$-quantile of the standard normal distribution, and $P(\log \hat{Z}_{onz})$ is a polynomial in $\log \hat{Z}_{onz}$. The role of $\widehat{\lambda}_{onz}$ is to correct for sample selection, while $P(\log \hat{Z}_{onz})$ addresses the extensive-margin bias by non-parametrically controlling for ϕ_{onz}.[31]

Data and Descriptive Statistics. To make the estimation sample consistent with our firm-level regressions, we construct our industry-level data by aggregating our firm-level data (described in Section 3 above) to the 6-digit HS level. We end up with information on export values, market shares and HHIs at the origin-destination-industry level for two exporting countries (France and China), 31 European destinations, and 1,864 industries for the year 2010. Data on business startup times and costs are sourced from the Worldbank's Doing Business Database.

Table 5 presents summary statistics on exporter HHIs and aggregate market shares of French and Chinese firms. It confirms that aggregate exports are concentrated among a small number of firms: the mean exporter HHI (which corresponds to HHI_{onz} in equation (20)) is 0.55; at the 90th percentile, a single firm accounts for the total market share of each country. Moreover, the mean aggregate market share of French and Chinese firms in each destination (s_{onz} in equation (20)) is around 10%; at the 90th percentile, that market share reaches 27%. Thus, in many markets, these exporters have substantial market power.

Estimation Results. Before presenting our results, it is worth pointing out that, as our industry-level data are constructed from firm-level data, we should expect to find estimates for the distance coefficient similar to those at the firm level. We first present results for

[31] A downside of the HMR method is that it does not account for potential heteroscedasticity. We have therefore also experimented with running PPML regressions, which however address neither sample selection nor the extensive-margin bias. The results from these regressions in combination with the Monte Carlo simulations in Section 5 below indicate that self-selection issues are much more severe than problems arising from potential heteroscedasticity, leading to a severe downward bias of the PPML estimates. We therefore do not report the PPML estimates.

Table 5: Summary Statistics for Industry-Level Market Shares and Exporter HHIs

	Exporter HHI	Destination Market Share
Mean	0.53	10%
5th pctile	0.07	0.01%
10th pctile	0.12	0.06%
Median	0.48	3%
90th pctile	1	27%
95th pctile	1	48%

Notes: Industry-level data. Table shows summary statistics on the distribution of exporter aggregate market share and HHI. The unit of observation is at the origin-destination-industry level. Sample for year 2010.

the pooled regressions where we constrain the distance coefficient to be the same across industries.

Table 6 reports OLS regression results with and without oligopoly correction, but without controlling for selection. The distance coefficient is -1.12 without, and, depending on the value of σ, in the range -1.23 to -1.29 with oligopoly correction. Thus, the oligopoly bias is still at work at the industry level, albeit of a smaller magnitude (around 10%) than at the firm level. The finding that these estimates are somewhat smaller than those at the firm level (around -1.5; see Table 3) suggests that the sample-selection bias is stronger than the extensive-margin bias.

Table 6: Industry-level Gravity Estimates without Controlling for Selection

	(1)	(2)	(3)	(4)
Method	OLS w/o corr	OLS w/ corr $\sigma = 4$	OLS w/ corr $\sigma = 5$	OLS w/corr $\sigma = 6$
log dist.	-1.128***	-1.227***	-1.260***	-1.293***
	(0.195)	(0.211)	(0.216)	(0.222)
Obs.	66,563	66,563	66,563	66,563
R^2	0.314	0.293	0.285	0.278
Or.-ind. FE	YES	YES	YES	YES
Dest.-ind. FE	YES	YES	YES	YES

Notes: Industry-level data. Standard errors clustered at destination level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

We now apply the HMR approach to correct for self-selection into exporting. Table 7 reports results from the first-step Probit estimation of the propensity to export.[32] As expected, the dummies for high business-startup cost and long business-startup time are negatively and

[32] We include 2-digit sector-origin and 2-digit sector-destination fixed effects, as using 6-digit industry-origin and industry-destination fixed effects is computationally infeasible with the Probit model. However, results

significantly associated with the propensity to export. Unsurprisingly, distance is also negatively related to export propensity, providing evidence for sample selection.

Table 7: First Step of HMR Procedure (Export Propensity)

	Export > 0
log dist.	-0.420*
	(0.234)
high startup cost	-1.340***
	(0.421)
long startup time	-2.102***
	(0.337)
Obs.	97,332
Sector-origin FE	YES
Sector-dest. FE	YES

Notes: Industry-level data. HMR step-one Probit regression of propensity to export. Standard errors clustered at destination level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 8 reports results from the second-step HMR regression. The specifications in columns (1)–(4) only include the inverse Mills ratio, and thus correct for sample selection but not for the extensive-margin effect. Columns (4)–(8) show the results from the full HMR procedure. These columns include a quadratic function of log $\hat{Z}_{onz}$[33] Throughout, the distance coefficient in the specifications with oligopoly correction is about 8–13% larger in magnitude than in those without. While the oligopoly-corrected estimated coefficients are still slightly smaller than those estimated at the firm level, they are larger than those from the OLS regression. This suggests that the sample-selection bias is slightly stronger than the extensive-margin bias. As expected, the specifications that only include the inverse Mills ratio yield the highest coefficient estimates, as they do not account for the extensive-margin bias.

Finally, we run the HMR procedure as before, but now separately for each 2-digit sector. Table 9 reports summary statistics across sectors on the distribution of the estimated coefficients and the magnitude of the oligopoly bias. While the oligopoly bias is relatively small in the median sector (around 10–17%, depending on the value of σ), it is substantial in a significant minority of sectors (around 47–68% at the 90th percentile). Moreover, the absolute oligopoly bias is positively correlated with the product of the average (across origins, destinations, and industries) Herfindahl index and the average exporting country's aggregate

using a linear probability model indicate hardly any changes in the point estimates when adding these more disaggregated fixed effects.

[33]We experimented with including higher-order terms, but this did not change the estimated distance coefficient, and the higher-order terms were not statistically significant.

Table 8: Industry-level Gravity Estimates—Controlling for Selection

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Meth.	Heck	Heck	Heck	Heck	HMR^2	HMR^2	HMR^2	HMR^2
	w/o	w/	w/	w/	w/o	w/	w/	w/
		$\sigma = 4$	$\sigma = 5$	$\sigma = 6$		$\sigma = 4$	$\sigma = 5$	$\sigma = 6$
log dist.	-1.189***	-1.293***	-1.327***	-1.362***	-1.151***	-1.250***	-1.284***	-1.317***
	(0.198)	(0.214)	(0.220)	(0.226)	(0.190)	(0.204)	(0.209)	(0.214)
inv. Mills	-0.121	-0.128	-0.130	-0.132	0.678***	0.768***	0.799***	0.829***
	(0.165)	(0.178)	(0.182)	(0.187)	(0.169)	(0.186)	(0.192)	(0.198)
$\log \hat{Z}$					0.907***	1.027***	1.067***	1.108***
					(0.265)	(0.285)	(0.293)	(0.300)
$\log \hat{Z}^2$					-0.103*	-0.119*	-0.125**	-0.130**
					(0.0565)	(0.0606)	(0.0621)	(0.0635)
Obs.	60,662	60,662	60,662	60,662	60,662	60,662	60,662	60,662
R^2	0.302	0.282	0.275	0.267	0.304	0.284	0.277	0.270
Ind-orig FE	YES	YES	YES	YES	YES	YES	YES	YES
Ind-dest FE	YES	YES	YES	YES	YES	YES	YES	YES

Notes: Industry-level data, pooled across industries, second step of HMR procedure. Standard errors clustered at destination level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

market share; the correlation coefficient is 0.16, suggesting a larger bias in more concentrated sectors.

Table 9: Industry-Level Gravity Estimates by 2-digit Sector

Median est coefficient	w/o corr	w/ corr $\sigma = 4$	w/corr $\sigma = 5$	w/corr $\sigma = 6$
log distance	-0.694	-0.819	-0.857	-0.893
abs. pct. bias (10th pctile)		3.2%	4.2%	5.2%
abs. pct. bias (median)		10.2%	13.7%	16.7%
abs. pct. bias (90th pctile)		68%	46.7%	52.3%

Notes: Industry-level data. Table shows summary statistics for the distribution of estimated coefficients from the HMR procedure by 2-digit HS sector.

5 Monte Carlo Simulations

In this section, we perform Monte Carlo simulations to evaluate the merits of our oligopoly correction terms. To this end, we develop and calibrate a model in which firms first self-select into export destinations and then compete in quantities. Using the calibrated model, we generate a Monte Carlo dataset to which we then apply our firm- and industry-level estimation procedures. We confirm that our oligopoly correction significantly improves the accuracy of our estimates.

Setup. The model is as described in Section 2, with $\lambda = 1$ (Cournot-Nash conduct). To avoid general-equilibrium effects (and, in the next section, to obtain a money-metric measure of social welfare), we assume that the representative consumer in each country has quasi-linear preferences:

$$U_n = q_{0n} + E_n \int_{z \in [0,1]} \log \left(\sum_{j \in \mathcal{J}_n(z)} a_{jn}(z)^{\frac{1}{\sigma}} q_{jn}(z)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} dz,$$

where q_{0n} denotes consumption of the outside good. For simplicity, we assume that parameters (such as the elasticity of substitution, or various technology parameters that are described in more detail below) do not vary across industries. Industries will still be heterogeneous due to different realizations of random variables such as productivity draws.

Each country has a fixed labor endowment. The outside good is freely traded and produced using only labor with a constant-returns-to-scale technology that is the same in all countries. We assume that parameters are such that it is produced in positive amount everywhere, so that its price is the same in every country. We further choose that good as the numeraire, which pins down the wage rate everywhere. In what follows, all costs should be understood as being incurred in terms of labor.

We focus on an industry $z \in [0,1]$ and drop the industry index to ease notation. We now put more structure on the distribution of cost and quality shocks, and on how firms make entry decisions into export destinations.

Recall from Section 2 that the cost for firm i of producing and selling q_{in} units in market n is $C_{in}(q_{in}) = c_{in}\tau_{in}q_{in}$. We decompose c_{in} log-linearly as $\log c_{in} = \varepsilon_i^c + \varepsilon_{in}^c$, where ε_i^c and ε_{in}^c are independent draws from normal distributions with mean zero and variance υ^2 and θ^2, respectively. The iceberg-type trade cost τ_{in} is set equal to 1 if firm i is based in country n and otherwise to $\tau_{on} \equiv T \times (\text{dist}_{on})^{\beta}$, where o denotes the country in which firm i is located, and T and β are parameters. Finally, we set a_{in} (the quality of product i in market n) equal to 1 for every i and n.[34]

A country-o firm that wants to sell in country $n \neq o$ must pay a fixed cost $f_{on} \equiv F \times \varphi_{on}^o \times \varphi_{on}^u$, where F is a parameter and φ_{on}^o and φ_{on}^u are i.i.d. draws from a standard log-normal distribution. The reason for this decomposition is that we will later assume that φ_{on}^o is observable to the econometrician whereas φ_{on}^u is not, so that φ_{on}^o can be used as an excluded first-stage variable when applying the HMR procedure. We set $f_{oo} = 0$ for every country o, so that a firm is always active in its home market.

We consider a two-stage game of complete information in which firms first simultaneously decide which markets to enter, and then compete in quantities in each market. Under oligopoly, this game is likely to have multiple subgame-perfect equilibria. If there were no fixed-cost heterogeneity, it would be possible to rank firms from highest to lowest (destination-specific) type and construct a subgame-perfect equilibrium in which high-type firms enter first. With fixed-cost heterogeneity (in addition to type heterogeneity), there is no such natural ranking of firms and constructing a subgame-perfect equilibrium is a non-trivial combinatorial problem. We therefore make the following simplifying behavioral assumption: When making entry decisions, firms believe that they will receive monopolistic-competition profits (given the set of firms that entered). We can then follow Spence (1976) and rank firms according to their survival coefficients, $(c_{in}\tau_{on})^{1-\sigma}/f_{on}$, in each market n. This pins down a natural "equilibrium" entry sequence in market n, in which firms with a higher survival

[34] Thus, using the notation of Section 2.2, we are setting

$$\varepsilon_n^c = \varepsilon_i^a = \varepsilon_n^a = \varepsilon_{in}^a = \varepsilon_i^\tau = \varepsilon_n^\tau = \varepsilon_{in}^\tau = 0.$$

The assumption that there is no destination-specific shock ($\varepsilon_n^c = \varepsilon_n^a = \varepsilon_n^\tau = 0$) is without loss of generality: as such shocks would affect all firms symmetrically, they would have no impact on equilibrium market shares and profits given CES demand. As for the firm and firm-destination quality and trade-cost shocks, we could alternatively assume that they are drawn i.i.d. from normal distributions and obtain an observationally equivalent model, as the resulting firm types would still be log-normally distributed.

coefficient enter first.

Calibration. We choose parameter values to generate a Monte Carlo dataset broadly similar to our firm-level dataset. We use the same set of countries as in the empirical implementation and take the bilateral distance matrix dist_{on} directly from the data. Market size in country n, E_n, is set equal to an amount proportional to that country's GDP in the data. We follow Chaney (2008) in assuming that the number of firms based in each country is proportional to its GDP. The proportionality coefficient is chosen so that the total number of firms is 220, which is similar to the number of firms in the average industry in our dataset. The elasticity of substitution σ is set to 5, as in our baseline empirical specification. Finally, we set $\beta = 0.38$, which is our baseline empirical estimate of the distance coefficient for $\sigma = 5$.

We still require values for the following four parameters: F, the intercept of the fixed-cost function; T, the intercept of the trade-cost function; υ, the standard deviation of firm baseline productivity draws; and θ, the standard deviation of firm-destination productivity shocks. We calibrate those parameters to match the following empirical moments (computed using the French and Chinese firm-level data): 1. the fraction of firm-destination-industry-year observations with zero trade flows (92%); 2. the mean (by destination-industry-year) aggregate combined market share of French and Chinese firms (13.9%); 3. the median (by origin-industry-year) 90/10 ratio of firm-level total exports (451); and 4. the median (by origin-destination-industry-year) 90/10 ratio of firm-destination exports (220).

The fact that each of the moments has a natural parameter counterpart gives rise to the following informal identification argument. Intuitively, we expect F to have a strong and negative effect on the first moment, T to have a strong and negative effect on the second moment, υ to have a strong and positive effect on the third moment, and θ to have a strong and positive effect on the fourth moment. In practice, we adjust the vector of parameters (F, T, υ, θ) to minimize the sum of the squared Davis-Haltiwanger deviations between theoretical and empirical moments.[35]

We approximate the theoretical moments using Monte Carlo integration. For each parameter vector, we perform 10 Monte Carlo runs.[36] For each run, we randomly draw vectors and matrices of firm-level baseline costs (ε_i^c), firm-destination cost shocks (ε_{in}^c), and fixed-cost

[35]The Davis-Haltiwanger deviation (Davis, Haltiwanger, and Schuh, 1996) is defined as the difference between the theoretical and empirical moments, divided by the arithmetic average of the theoretical and empirical moments. This residual converges to the percentage deviation when the theoretical moment tends to the empirical moment. The advantage of using this residual for our calibration procedure is that, in contrast to the percentage deviation, it always remains bounded and gives rise to symmetric punishments for positive and negative deviations.

[36]Note that, while the number of Monte Carlo runs is small, each run generates data for about 100 firms and 33 destinations, so that there is relatively little variation in theoretical moments across runs. Increasing the number of runs beyond 10 would only make a small difference, but would substantially increase computational requirements.

shocks (φ^o_{on}) and (φ^u_{on}). For each destination within a run, we then compute the equilibrium of the entry stage and, using a variant of Nocke and Schutz (2018)'s nested fixed-point algorithm, the equilibrium of the quantity-setting stage. Having done that for all ten runs, we compute arithmetic means (for moments 1 and 2) and medians (for moments 3 and 4) to obtain Monte Carlo approximations to our theoretical moments.

Our calibration algorithm converges to $F = 3.62 \times 10^{-9}$ (times total world expenditures in the industry, which we normalized to unity), $T = 0.827$, $\upsilon = 0.394$, and $\theta = 1.23$. We obtain nearly perfect matches for the second, third, and fourth moments (0.140, 449, and 220, respectively, vs. 0.139, 451, and 220 in the data), and we slightly under-predict the fraction of zeros in the firm-level export matrix (82.8% vs. 92% in the data). The resulting sum of squared deviations is 0.011.

Data generation and results. Using the calibrated parameters, we generate the Monte Carlo dataset. We perform 200 Monte Carlo runs. Each run features different realizations of productivity and fixed-cost shocks, and can thus be thought of as a different industry or a different time period. For each run, we compute the equilibrium of the entry model and of the quantity-setting game in all markets, and we store firm-level sales and market shares, origin, destination, firm, and run indicators, and bilateral distance and observable fixed-cost shocks. To make the dataset comparable to the one used in our empirical applications, we keep observations only for firms based in the countries corresponding to France and China. We thus obtain a firm-level dataset, which we also aggregate up to construct an industry-level dataset. On these, we run our firm- and industry-level regressions to evaluate the performance of our oligopoly correction terms.

Table 10 reports the results from our firm-level regressions. As the OLS estimates are strongly biased towards zero, consistent with the heteroscedasticity bias discussed in Section 3, we focus in that table on the PPML estimates.[37] The specification without oligopoly correction (column (1)) significantly underestimates the absolute value of the distance coefficient. The specification with oligopoly correction (column (2)) delivers an estimate that is very similar to the true distance coefficient (-1.52). Interestingly, the (biased) PPML estimate without oligopoly correction is very close to the empirical estimate in Table 3 (-0.851 vs. -0.874). Thus, our Monte Carlo dataset generates an oligopoly bias almost identical in size to the one obtained in our empirical analysis.[38]

Table 11 reports the estimation results from industry-level regressions. As the PPML

[37]The OLS estimate is -0.454 without oligopoly correction and -0.472 with oligopoly correction. Similarly, the PPML estimate when dropping the observations with zero trade flows is -0.717 without oligopoly correction and -1.024 with oligopoly correction. See Table F in Online Appendix D.

[38]As shown in Table G in Online Appendix D, very similar results obtain when using (i) only the top French and top Chinese exporter and (ii) the top-5 French and top-5 Chinese exporters.

Table 10: Monte Carlo: Firm-Level Results

Method	(1) PPML w/o corr	(2) PPML w/ corr $\sigma = 5$
log dist.	-0.851***	-1.403***
	(0.135)	(0.365)
Obs.	30,198	30,198
(Pseudo) R^2	0.21	0.74
Firm-run FE	YES	YES
Destination-run FE	YES	YES

Notes: Monte Carlo dataset, firm-level data, pooled across Monte Carlo runs. Results for top-3 exporters, without and with oligopoly correction. Standard errors clustered at the destination level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. True log-distance coefficient is -1.52.

estimates are strongly biased toward zero, presumably due to selection effects, we focus in that table on OLS and HMR estimates, with and without oligopoly correction.[39] In all specifications, our oligopoly correction improves the accuracy of the distance-coefficient estimate. Specifications that account for self-selection (columns (3)–(6)), when combined with our correction term, deliver estimates that are very close to the true value of -1.52.

6 Counterfactual Simulations

In this section, we turn to the welfare effects of a trade liberalization, and quantitatively assess the importance of accounting for oligopolistic behavior. We do so by calibrating two versions of the model of Section 5: the oligopoly ('*oli*') version, in which $\lambda = 1$ (Cournot-Nash conduct) and the distance coefficient is set equal to our baseline empirical estimate *with* the oligopoly correction term; and the monopolistic competition ('*mc*') version in which $\lambda = 0$ (monopolistic-competition conduct) and the distance coefficient is our baseline estimate *without* the correction term. We calibrate both versions by matching the same moments in the data, and then use them to simulate a 10% trade-cost reduction and compute the induced welfare effects.

Setup. The setup is as described in Section 5, with the following amendments. First, at the entry stage of the *oli* version, firms now correctly expect to earn oligopoly profits. As discussed above, in the *oli* version such correct conjectures make it infeasible to solve for a subgame-perfect equilibrium under fixed-cost heterogeneity. We thus, second, assume that

[39] The PPML estimate without oligopoly correction is -0.849. Correcting for oligopoly forces results in a larger estimated distance coefficient of -1.04.

Table 11: Monte Carlo: Industry-Level Results

Method	(1) OLS w/o	(2) OLS w/ $\sigma = 5$	(3) Heck w/o	(4) Heck w/ $\sigma = 5$	(5) HMR w/o	(6) HMR w/ $\sigma = 5$
log dist.	-1.272***	-1.383***	-1.282***	-1.399***	-1.298***	-1.418***
	(0.0658)	(0.0728)	(0.0957)	(0.108)	(0.0990)	(0.111)
inv. Mills			0.795	0.909	-3.482	-3.584
			(0.523)	(0.610)	(2.295)	(2.573)
$\log \hat{Z}$					-4.227	-4.386
					(2.593)	(2.897)
$\log \hat{Z}^2$					0.827	0.851
					(0.544)	(0.602)
Obs.	11,094	11,094	8,296	8,296	8,296	8,296
R^2	0.64	0.61	0.64	0.62	0.64	0.62
Or.-run FE	YES	YES	YES	YES	YES	YES
Dest.-run FE	YES	YES	YES	YES	YES	YES

Notes: Monte Carlo dataset, industry-level data, pooled across Monte Carlo runs. Standard errors clustered at the destination level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. True log-distance coefficient is -1.52.

the fixed export cost, $f > 0$, does not vary across origin-destination pairs. This allows us to rank firms from highest to lowest type and construct an equilibrium of the entry game in which firms with a higher type enter first. Finally, we increase the number of firms based in each country by 1 to ensure that each market is always served by at least two firms, so that consumer surplus is always finite.[40]

Calibration. The set of countries, the bilateral distance matrix, each country's market size, the coefficient that determines the number of firms in each country, and the elasticity of substitution are as in Section 5. The distance coefficient, β, is set to 0.38 in the *oli* version and to 0.22 in the *mc* version, which corresponds to our baseline empirical estimates with and without oligopoly correction (for $\sigma = 5$); see columns (3) and (1) in Table 3. The remaining parameters (f, T, υ, and θ) are chosen to match the same moments as in the previous section. The theoretical moments are again approximated using Monte Carlo integration with 10 iterations.

For the *oli* version, the calibration algorithm converges to $f = 1.36 \times 10^{-8}$, $T = 0.379$, $\upsilon = 0.406$, and $\theta = 1.15$. The calibrated model does a very good job of matching the 90/10 dispersion moments (218 and 456 for firm-destination and firm-level exports, respectively, vs. 220 and 451 in the data) and the mean aggregate share of French and Chinese firms (13.6%

[40]Under CES demand, a monopolist would set an infinite price, resulting in consumer surplus being equal to minus infinity.

vs. 13.9% in the data) but tends to under-predict the fraction of zeros in the firm-level export matrix (79.8% vs. 92% in the data). This results in a sum of squared deviations of 0.0207. The fit of the *mc* calibration is almost as good, with a sum of squared residuals of 0.0262. The calibrated model continues to provide a good match for the 90/10 and aggregate-share moments (221, 455, and 14.2%, respectively) but still under-predicts the fraction of zeros (78.4%). The values of the productivity parameters υ and θ are close to the *oli* calibration ($\upsilon = 0.257$ and $\theta = 1.15$). Productivities are slightly less dispersed in the *mc* calibration, which is intuitive, as the sales distribution tends to be more compressed under oligopoly due to incomplete passthrough. As profits tend to be lower under monopolistic competition, the calibrated fixed cost ($f = 5.67 \times 10^{-9}$) is lower than in the *oli* calibration. Finally, the fact that the distance coefficient β is significantly lower in the *mc* calibration mechanically reduces trade costs to all destinations. This results in the intercept of the trade cost function ($T = 1.62$) being higher than in the *oli* calibration, so as not to overpredict the exports of French and Chinese firms.

Computing social welfare. Plugging country n's budget constraint into the representative consumer's utility function, we obtain an expression for social welfare in that country (up to an additive constant):

$$W_n = \int_{z\in[0,1]} \left(E_n \left[\frac{\sigma}{\sigma-1} \log \left(\sum_{j\in\mathcal{J}_n(z)} q_{jn}(z)^{\frac{\sigma-1}{\sigma}} \right) - 1 \right] + \Pi_n(z) \right) dz, \qquad (22)$$

where $\Pi_n(z)$ represents the total profits made by firms based in country n.[41] To report the values of our welfare measures in U.S. dollars, we set E_n equal to country n's GDP share in our dataset multiplied by the value added in manufacturing, added up over all 33 countries. As in Section 5, the integral in equation (22) is approximated using Monte Carlo integration with 200 iterations (i.e., 200 industries).

Results. We simulate the equilibrium effects of a 10% reduction in variable trade costs using the *oli* and *mc* versions of the model. Figure 1 reports the resulting changes in social welfare per capita.[42] According to our simulations, the welfare gains from a trade liberalization are substantially higher under oligopoly, with the average European individual experiencing a utility gain of USD 319 in the *oli* version and USD 181 in the *mc* version. A similar picture

[41] We are thus assuming that firms are owned by the residents of their country of origin. The results are very similar when assuming instead that consumers own an internationally diversified portfolio.

[42] To improve the figure's readability, we have dropped two outliers, Iceland and Luxemburg, for which the gains from trade under oligopoly significantly exceed USD 1,500 per capita. We have also dropped China, which, in this model, benefits very little from trade liberalization, due to it being a remote market to which very few firms find it profitable to export.

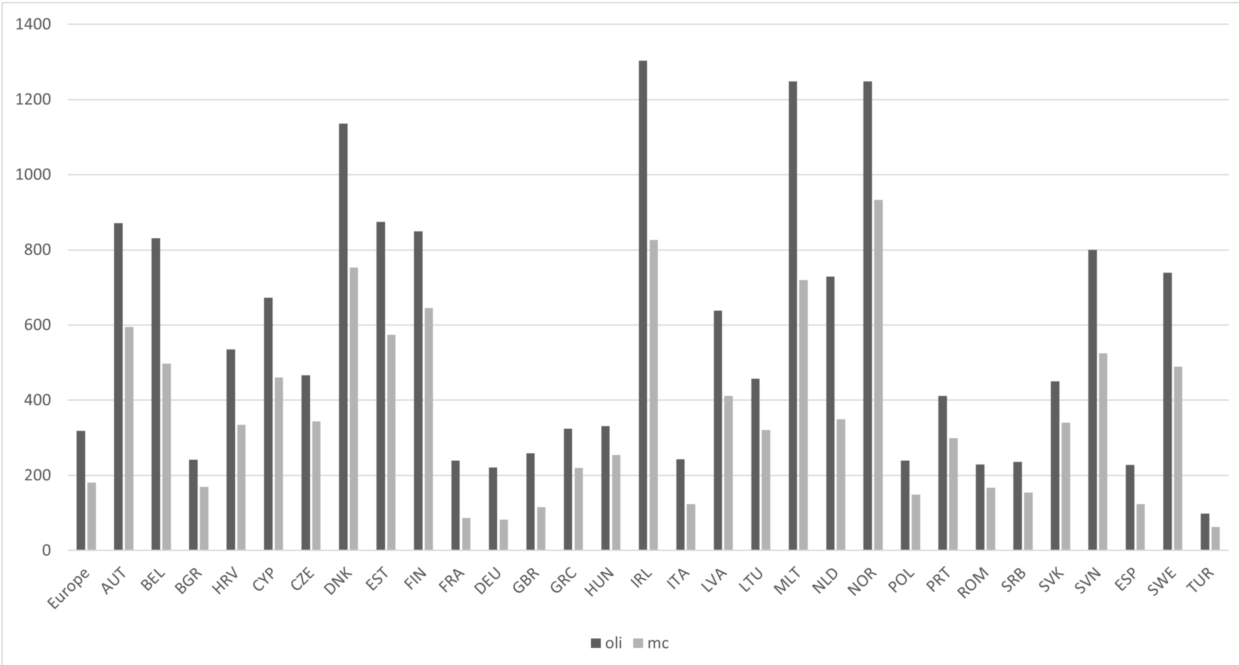

Figure 1: Welfare Effects of a 10% Trade Cost Reduction (in USD per Capita)

Notes: Figure shows the effects of a 10% trade cost reduction on per-capita welfare (measured in USD), by country, under oligopoly (*oli*) and monopolistic competition (*mc*).

emerges when looking at individual European countries, with most countries experiencing gains from trade that are at least 30% higher in the *oli* calibration than in the *mc* calibration. In large, central countries such as France or Germany, the gains from trade under oligopoly are almost three times as high as in the *mc* calibration.

To better understand what drives the difference between the *oli* and the *mc* predictions, we decompose the welfare effects of the trade liberalization into: 1. a trade-cost component (the marginal costs of all exporters decrease by 10%, holding fixed all markups and the set of exporters); 2. a domestic-markups component (due to increased competitive pressure, domestic firms lower their markups); 3. a foreign-markups component (exporters, whose market shares have increased, raise their markups); and 4. an extensive-margin component (the set of exporters adjust).

We now report on the magnitude of these components for European social welfare per capita; the general picture is similar when looking at individual European countries. In our simulations, the extensive-margin component is negligible under both oligopoly and monopolistic competition. The domestic-markups component raises per-capita welfare by USD 67 in the *oli* version, while the foreign-markups component lowers it by USD 20; both components are of course inoperative under monopolistic competition. Finally, the trade-cost component raises per-capita welfare by USD 272 in the *oli* version, and by USD 181 in the *mc* version.

Thus, around two-thirds of the gap between the gains from trade under oligopoly and monopolistic competition can be attributed to the fact that the *oli* calibration results in different trade cost parameters, with the remaining third being explained by markup adjustments.[43] This highlights the importance of obtaining a reliable estimate of the distance coefficient β.

7 Conclusion

We have developed an international trade model with CES preferences and heterogeneous firms that are granular and thus behave oligopolistically. In this model, we have derived a gravity equation of international trade flows, both at the firm and industry level. We have shown that the standard approach to gravity estimation suffers from an omitted variable bias when firms behave oligopolistically. We have proposed methods to purge the observed trade flows from market-power effects and thus obtain consistent estimates of gravity parameters. Using French and Chinese export data and Monte Carlo simulations, we have shown that accounting for oligopoly is quantitatively important. When estimating gravity at the firm level, the elasticity of trade flows with respect to distance is more than 70% larger when correcting for market power. While the magnitude of the oligopoly bias is smaller when estimating gravity at the industry level, it is still substantial in a significant minority of industries, namely in those in which exports tend to be highly concentrated. In a calibrated version of our model, the welfare gains from a trade liberalization are almost twice as large under oligopoly as under monopolistic competition. These findings reinforce the view that market power effects matter in international trade.

Appendix

A Proofs and Derivations

A.1 Proof of Proposition 1

Proof. To complete the proof of the proposition, we need to: (a) show that the function S is well defined, and study its monotonicity properties and its limits; (b) show that the equilibrium condition (12) has a unique solution; (c) show that, at $\lambda = 1$, the first-order

[43] Our results thus contrast with those of Arkolakis, Costinot, Donaldson, and Rodriguez-Clare (2019) in two dimensions. First, while they find that the welfare gains from trade liberalization are lower under variable markups, we find larger gains under oligopoly than under monopolistic competition with constant markups. Second, while they report that the (negative) foreign-markups component more than outweighs the (positive) domestic-markups component, we obtain the opposite.

conditions of profit maximization are sufficient for global optimality. We do so below. In the following, we drop the destination index (n) to ease notation.

(a). As the right-hand of equation (11) is strictly increasing in s_i whereas the left-hand side is non-increasing in s_i, this equation has at most one solution. As s_i tends to 0, the left-hand side of that equation tends to 1, whereas the right-hand side tends to 0. As s_i tends to ∞, the left-hand side tends to 1 or $-\infty$, and the right-hand side tends to $+\infty$. The equation therefore has a unique solution, $S(T_i/H, \lambda) \in (0, 1/\lambda)$, where $1/\lambda \equiv \infty$ when $\lambda = 0$. It is easily checked that $S(\cdot, \cdot)$ is strictly increasing in its first argument and strictly decreasing in its second argument. By monotonicity, $S(\cdot, \lambda)$ has limits at 0 and ∞. Clearly, those limits are equal to 0 and $1/\lambda$, respectively.

(b). The results in the previous paragraph imply that the left-hand side of equation (12) is strictly decreasing in H, and has limits 0 and $|\mathcal{J}|/\lambda$ as H tends to ∞ and 0, respectively. It follows that equation (12) has a unique solution, $H^*(\lambda)$.

(c). Rewriting equation (2) with $\lambda = 1$ yields:

$$\frac{\partial \pi_i}{\partial q_i} = \frac{\sigma - 1}{\sigma} E \frac{a_i^{\frac{1}{\sigma}} q_i^{-\frac{1}{\sigma}}}{\sum_{j \in \mathcal{J}} a_j^{\frac{1}{\sigma}} q_j^{\frac{\sigma-1}{\sigma}}} \left(1 - \frac{a_i^{\frac{1}{\sigma}} q_i^{\frac{\sigma-1}{\sigma}}}{\sum_{j \in \mathcal{J}} a_j^{\frac{1}{\sigma}} q_j^{\frac{\sigma-1}{\sigma}}} \right) - c_i \tau_i,$$

which is decreasing in q_i. Hence, π_i is concave in q_i, and so firm i's first-order condition is sufficient for global optimality. □

A.2 Proof of Proposition 2

Proof. To apply Taylor's theorem, we require the value of $s_{on}^{*\prime}(0)$. We thus need to compute the partial derivatives of $S(\cdot, \cdot)$ at $\lambda = 0$ and $H_n^{*\prime}(0)$. Differentiating equation (11) with respect to s_{in}, λ, and $t_{in} \equiv T_{in}/H_n$ at $\lambda = 0$ yields

$$-s_{in} d\lambda = \frac{1}{\sigma - 1} \frac{ds_{in}}{s_{in}} - \frac{\sigma}{\sigma - 1} \frac{dt_{in}}{t_{in}}.$$

It follows that[44]

$$t_{in} \partial_1 \log S(t_{in}, 0) = \sigma \quad \text{and} \quad \partial_2 \log S(t_{in}, 0) = -(\sigma - 1) S(t_{in}, 0).$$

[44] Notation: $\partial_k S$ is the partial derivative of S with respect to its kth argument.

Next, we differentiate equation (12) with respect to λ and H_n:

$$\sum_{j\in\mathcal{J}_n}\left[-\frac{T_{jn}}{H_n}\partial_1 S\left(\frac{T_{jn}}{H_n},\lambda\right)\frac{dH_n}{H_n}+\partial_2 S\left(\frac{T_{jn}}{H_n},\lambda\right)d\lambda\right]=0.$$

Setting $\lambda=0$ and plugging in the values of the partial derivatives of S, we obtain:

$$\sum_{j\in\mathcal{J}_n}\left[\sigma s^*_{jn}(0)\frac{dH_n}{H_n}+(\sigma-1)\left(s^*_{jn}(0)\right)^2 d\lambda\right]=0.$$

Making use of the definition of $\text{HHI}_n(0)$ and of the fact that market shares add up to one, we obtain:

$$\frac{H_n^{*\prime}(0)}{H_n^*(0)}=-\frac{\sigma-1}{\sigma}\,\text{HHI}_n(0).$$

We can now compute $s^{*\prime}_{in}(0)$:

$$\begin{aligned}
s^{*\prime}_{in}(0) &= \frac{\partial}{\partial\lambda}S\left(\frac{T_{in}}{H_n^*(\lambda)},\lambda\right)\bigg|_{\lambda=0}\\
&= -\frac{T_{in}}{H_n^*(0)}\partial_1 S\left(\frac{T_{in}}{H_n^*(0)},0\right)\frac{H_n^{*\prime}(0)}{H_n^*(0)}+\partial_2 S\left(\frac{T_{in}}{H_n^*(0)},0\right)\\
&= (\sigma-1)\left[s^*_{in}(0)\,\text{HHI}_n(0)-(s^*_{in}(0))^2\right].
\end{aligned}$$

It follows that

$$\begin{aligned}
\frac{s^{*\prime}_{on}(0)}{s^*_{on}(0)} &= (\sigma-1)\frac{1}{s^*_{on}(0)}\sum_{j\in\mathcal{J}_{on}}\left[s^*_{jn}(0)\,\text{HHI}_n(0)-\left(s^*_{jn}(0)\right)^2\right]\\
&= (\sigma-1)\left[\text{HHI}_n(0)-s^*_{on}(0)\sum_{j\in\mathcal{J}_{on}}\left(\frac{s^*_{jn}(0)}{s^*_{on}(0)}\right)^2\right]\\
&= (\sigma-1)\left[\text{HHI}_n(0)-s^*_{on}(0)\,\text{HHI}_{on}(0)\right].
\end{aligned}$$

Applying Taylor's theorem at the first order in the neighborhood of $\lambda=0$ yields:

$$\begin{aligned}
\log s^*_{on}(\lambda) &= \log s^*_{on}(0)+\frac{d}{d\lambda}\log s^*_{on}(\lambda)\bigg|_{\lambda=0}\lambda+o(\lambda)\\
&= \log s^*_{on}(0)+(\sigma-1)\left[\text{HHI}_n(0)-s^*_{on}(0)\,\text{HHI}_{on}(0)\right]\lambda+o(\lambda)\\
&= \log s^*_{on}(0)+(\sigma-1)\left[\text{HHI}_n(\lambda)-s^*_{on}(\lambda)\,\text{HHI}_{on}(\lambda)\right]\lambda+o(\lambda),
\end{aligned}$$

where the last line follows as $\text{HHI}_n(\lambda)-\text{HHI}_n(0)$ and $s^*_{on}(\lambda)\,\text{HHI}_{on}(\lambda)-s^*_{on}(0)\,\text{HHI}_{on}(0)$ are at most first order. □

A.3 Further Theoretical Results

Further results for Section 2.2. We begin by deriving equation (9) in the main text using the aggregative games approach presented in Section 2.3. We are interested in the partial derivatives of $\log r_{in}$ with respect to the bilateral variables holding fixed the characteristics of the destination market:[45]

$$\begin{aligned}
\nabla_{X_{in}} \log r_{in} &= \frac{\partial \log T_{in}}{\partial X_{in}} \frac{\partial}{\partial \log T_{in}} (\log s_{in} + \log E_n) \\
&= -\frac{\sigma - 1}{\sigma} \frac{\partial \log \tau_{in}}{\partial X_{in}} \frac{\partial \log S(T_{in}, H_n)}{\partial \log T_{in}} \\
&= -\frac{\sigma - 1}{\sigma} \beta \frac{\partial \log S(T_{in}, H_n)}{\partial \log T_{in}},
\end{aligned}$$

where we have used the definition of T_{in} and S (equations (10) and (11)) to obtain the second line and the parameterization of $\log \tau_{in}$ to obtain the third line. The partial derivative $\partial S(T_{in}, H_n)/\partial T_{in}$ can be obtained by applying the implicit function theorem to equation (11):

$$\frac{\partial \log S(T_{in}, H_n)}{\partial \log T_{in}} = \frac{\sigma}{1 + (\sigma - 1)\frac{S(T_{in}, H_n)}{1 - S(T_{in}, H_n)}}.$$

Combining this with the above expression and simplifying, we obtain equation (9) in the main text.

Further results for Section 2.3. Finally, we derive equation (15) in the main text. We are interested in the partial derivatives of $\log r_{on}$ with respect to the bilateral variables holding fixed the characteristics of the destination market:

$$\begin{aligned}
\nabla_{X_{on}} \log r_{on} &= \frac{1}{s_{on}} \sum_{j \in \mathcal{J}_{on}} s_{jn} \frac{\partial \log s_{jn}}{\partial X_{on}} \\
&= \beta(1 - \sigma) \frac{1}{s_{on}} \sum_{j \in \mathcal{J}_{on}} s_{jn} \frac{1}{1 + (\sigma - 1)\frac{s_{jn}}{1 - s_{jn}}} \\
&\simeq \beta(1 - \sigma) \frac{1}{s_{on}} \sum_{j \in \mathcal{J}_{on}} s_{jn} \left(1 - (\sigma - 1) s_{jn}\right) \\
&= \beta(1 - \sigma) \left(1 - (\sigma - 1) s_{on} \, \mathrm{HHI}_{on}\right),
\end{aligned}$$

where we have used equation (9) to obtain the second line and a Taylor approximation around $s_{jn} \simeq 0$ to obtain the third line. Equation (15) follows.

[45] If X_{in} is multidimensional, $\partial/\partial X_{in}$ should be understood as $\nabla_{X_{in}}$.

B Estimating σ Using Tariff Data

In this section, we show how the elasticity of substitution, σ, can be jointly estimated with the coefficients on bilateral variables, β, if the data feature sufficient variation in tariffs. Following the same steps as in Section 2.2, we obtain the firm-level gravity equation

$$\log r_{in} = \xi_n + \zeta_i + X_{in}\beta(1-\sigma) + (1-\sigma)\log(1+t_{in}) + (\sigma-1)\log(1-\mu_{in}) + \varepsilon_{in}, \tag{23}$$

where t_{in} is the ad-valorem tariff paid by firm i in destination n.

Using the formula for the equilibrium markup (equation (4)), absorbing the constant term $(\sigma-1)\log((\sigma-1)/\sigma)$ into the fixed effects, and introducing the dummy variables

$$\check{D}^j_{in} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \quad \text{and } \tilde{D}^m_{in} = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{otherwise} \end{cases},$$

we can rewrite equation (23) as

$$\begin{aligned} \log r_{in} &= \sum_m \xi_m \tilde{D}^m_{in} + \sum_j \zeta_j \check{D}^j_{in} + X_{in}\beta(1-\sigma) + (1-\sigma)\log(1+t_{in}) + (\sigma-1)\log(1-s_{in}) + \varepsilon_{in} \\ &= \tilde{D}_{in}\xi + \check{D}_{in}\zeta + X_{in}\beta(1-\sigma) + (1-\sigma)\log(1+t_{in}) + (\sigma-1)\log(1-s_{in}) + \varepsilon_{in}, \end{aligned}$$

where ξ and ζ are the column vectors of destination and firm fixed effects, respectively, and $\tilde{D}_{in}$ and $\check{D}_{in}$ are row vectors of destination and firm dummy variables.

Using the above gravity equation to eliminate ε_{in} from the orthogonality conditions $\mathbb{E}(\tilde{D}_{in}\varepsilon_{in}) = 0$, $\mathbb{E}(\check{D}_{in}\varepsilon_{in}) = 0$, $\mathbb{E}(X_{in}\varepsilon_{in}) = 0$, and $\mathbb{E}(\log(1+t_{in})\varepsilon_{in}) = 0$, we obtain the moment conditions

$$\begin{aligned} \mathbb{E}\Big(\Big[&\log r_{in} - \tilde{D}_{in}\xi - \check{D}_{in}\zeta - X_{in}\beta(1-\sigma) \\ &-(1-\sigma)\log(1+t_{in}) - (\sigma-1)\log(1-s_{in})\Big]\Big(\tilde{D}_{in}, \check{D}_{in}, X_{in}, \log(1+t_{in})\Big)\Big) = 0. \end{aligned}$$

Using these moment conditions, the parameters of interest, β and σ, can be consistently estimated by GMM.[46]

[46] As in the main text, potential selection issues can be mitigated by focusing on top exporters. To address concerns of heteroscedasticity, the GMM objective function can be constructed by writing the moment conditions in levels rather than logs, mirroring the PPML estimator in the main text.

References

Allen, T., C. Arkolakis, and Y. Takahashi (2020): "Universal Gravity," *Journal of Political Economy*, 128(2), 393–433.

Anderson, J. (1979): "A theoretical foundation for the gravity equation," *American Economic Review*, 69(1), 106–116.

Anderson, J., and E. van Wincoop (2003): "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review*, 93(1), 170–192.

Anderson, S. P., N. Erkal, and D. Piccinin (2020): "Aggregative games and oligopoly theory: short-run and long-run analysis," *RAND Journal of Economics*, 51(2), 470–495.

Angrist, J. D., and J.-S. Pischke (2009): *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Arkolakis, C., A. Costinot, D. Donaldson, and A. Rodriguez-Clare (2019): "The Elusive Pro-Competitive Effects of Trade," *The Review of Economic Studies*, 86(1), 46–80.

Arkolakis, C., A. Costinot, and A. Rodriguez-Clare (2012): "New Trade Models, Same Old Gains?," *American Economic Review*, 102(1), 94–130.

Atkeson, A., and A. Burstein (2008): "Pricing-to-Market, Trade Costs, and International Relative Prices," *American Economic Review*, 98(5), 1998–2031.

Atkin, D., and D. Donaldson (2015): "Who's Getting Globalized? The Size and Implications of Intra-national Trade Costs," NBER Working Papers 21439, National Bureau of Economic Research, Inc.

Baier, S. L., and J. H. Bergstrand (2007): "Do free trade agreements actually increase members' international trade?," *Journal of International Economics*, 71(1), 72–95.

Bas, M., T. Mayer, and M. Thoenig (2017): "From micro to macro: Demand, supply, and heterogeneity in the trade elasticity," *Journal of International Economics*, 108(C), 1–19.

Brander, J. A. (1995): "Chapter 27: Strategic Trade Policy," in *Handbook of International Economics*, ed. by G. Grossman, and K. Rogoff, vol. 3, pp. 1395 – 1455. Elsevier.

Breinlich, H., H. Fadinger, V. Nocke, and N. Schutz (2023): "Gravity with Granularity," CEPR Discussion Papers 15374 (v.2), C.E.P.R. Discussion Papers.

Breinlich, H., V. Nocke, and N. Schutz (2020): "Merger Policy in a Quantitative Model of International Trade," *The Economic Journal*, 130(626), 393–421.

Bresnahan, T. F. (1989): "Chapter 17: Empirical Studies of Industries with Market

Power," in *Handbook of Industrial Organization*, ed. by R. Schmalensee, and R. Willig, vol. 2, pp. 1011–1057. Elsevier.

Broda, C., and D. E. Weinstein (2006): "Globalization and the Gains from Variety," *Quarterly Journal of Economics*, 121(2), 541–588.

Burstein, A., and G. Gopinath (2014): "International Prices and Exchange Rates," in *Handbook of International Economics*, vol. 4, chap. Chapter 7, pp. 391–451. Elsevier.

Chaney, T. (2008): "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, 98(4), 1707–1721.

Correia, S., P. Guimaraes, and T. Zylkin (2019): "PPMLHDFE: Stata module for Poisson pseudo-likelihood regression with multiple levels of fixed effects," Statistical Software Components, Boston College Department of Economics.

Davis, S. J., J. Haltiwanger, and S. Schuh (1996): *Job Creation and Destruction*. Cambridge, MA: MIT Press.

Eaton, J., and S. Kortum (2002): "Technology, Geography, And Trade," *Econometrica*, 70(5), 1741–1779.

Eckel, C., and P. Neary (2010): "Multi-Product Firms and Flexible Manufacturing in the Global Economy," *Review of Economic Studies*, 77(1), 188–217.

Edmond, C., V. Midrigan, and D. Y. Xu (2015): "Competition, Markups, and the Gains from International Trade," *American Economic Review*, 105(10), 3183–3221.

Feenstra, R. C. (1989): "Symmetric pass-through of tariffs and exchange rates under imperfect competition: An empirical test," *Journal of International Economics*, 27(1-2), 25–45.

——— (1994): "New Product Varieties and the Measurement of International Prices.," *American Economic Review*, 84(1), 157–177.

Ganapati, S., J. S. Shapiro, and R. Walker (2020): "Energy Cost Pass-Through in US Manufacturing: Estimates and Implications for Carbon Taxes," *American Economic Journal: Applied Economics*, 12(2), 303–342.

Gaubert, C., and O. Itskhoki (2021): "Granular Comparative Advantage," *Journal of Political Economy*, 121(3).

Genakos, C., and M. Pagliero (2022): "Competition and Pass-Through: Evidence from Isolated Markets," *American Economic Journal: Applied Economics*, 14(4), 35–57.

Harrigan, J. (1996): "Openness to trade in manufactures in the OECD," *Journal of International Economics*, 40(1-2), 23–39.

Head, K., and T. Mayer (2014): "Gravity Equations: Workhorse, Toolkit, and Cookbook," in *Handbook of International Economics*, vol. 4, chap. Chapter 3, pp. 131–195. Elsevier.

——— (2023): "Poor Substitutes? Counterfactual methods in IO and Trade compared," *The Review of Economics and Statistics*, forthcoming.

Heid, B., and F. Stähler (2024): "Structural gravity and the gains from trade under imperfect competition: Quantifying the effects of the European Single Market," *Economic Modelling*, 131(C).

Helpman, E., M. Melitz, and Y. Rubinstein (2008): "Estimating Trade Flows: Trading Partners and Trading Volumes," *Quarterly Journal of Economics*, 123(2), 441–487.

Khandelwal, A. K., P. K. Schott, and S.-J. Wei (2013): "Trade Liberalization and Embedded Institutional Reform: Evidence from Chinese Exporters," *American Economic Review*, 103(6), 2169–2195.

Melitz, M., and G. Ottaviano (2008): "Market Size, Trade, and Productivity," *Review of Economic Studies*, 75, 295–316.

Melitz, M., and S. Redding (2015): "New Trade Models, New Welfare Implications," *American Economic Review*, 105(3), 1105–46.

Nakamura, E., and D. Zerom (2010): "Accounting for Incomplete Pass-Through," *Review of Economic Studies*, 77(3), 1192–1230.

Neary, J. P. (2003): "Presidential Address: Globalization and Market Structure," *Journal of the European Economic Association*, 1(2-3), 245–271.

Nocke, V., and N. Schutz (2018): "Multiproduct-Firm Oligopoly: An Aggregative Games Approach," *Econometrica*, 86(2), 523–557.

——— (2024): "An Aggregative Games Approach to Merger Analysis in Multiproduct-Firm Oligopoly," *RAND Journal of Economics,* forthcoming.

Parenti, M. (2018): "Large and small firms in a global market: David vs. Goliath," *Journal of International Economics*, 110, 103–118.

Santos Silva, J. M. C., and S. Tenreyro (2006): "The Log of Gravity," *The Review of Economics and Statistics*, 88(4), 641–658.

Spence, A. M. (1976): "Product Selection, Fixed Costs, and Monopolistic Competition," *Review of Economic Studies*, 43(2), 217–235.

Tinbergen, J. (1962): *Shaping the World Economy: Suggestions for an International Economic Policy.* Twentieth Century Fund, New-York.

Van Beveren, I., A. B. Bernard, and H. Vandenbussche (2012): "Concording EU

Trade and Production Data over Time," NBER Working Papers 18604, National Bureau of Economic Research, Inc.

Yitzhaki, S. (1996): "On Using Linear Regressions in Welfare Economics," *Journal of Business & Economic Statistics*, 14(4), 478–486.

Online Appendix to:
Gravity with Granularity

Holger Breinlich* Harald Fadinger† Volker Nocke‡ Nicolas Schutz§

September 12, 2024

C Price Competition

C.1 Theoretical Results

Under price competition, the profit of firm i when selling in destination n is:

$$\pi_{in} = (p_{in} - \tau_{in}c_{in})a_{in}p_{in}^{-\sigma}P_n^{\sigma-1}\alpha_n E_n,$$

where we have dropped the industry index z for ease of notation.

The degree of strategic interactions between firms continues to be governed by the conduct parameter $\lambda \in [0, 1]$: When firm i increases its price by an infinitesimal amount, it perceives the induced effect on P_n to be equal to $\lambda \partial P_n/\partial p_{in}$. It is still the case that monopolistic competition arises when $\lambda = 0$, whereas Bertrand competition arises when $\lambda = 1$. The first-order condition of profit maximization of firm i in destination n is given by

$$\begin{aligned} 0 = \frac{\partial \pi_{in}}{\partial p_{in}} &= a_{in}p_{in}^{-\sigma}P_n^{\sigma-1}\alpha_n E_n + (p_{in} - \tau_{in}c_{in})\left[-\frac{\sigma}{p_{in}} + \frac{\sigma-1}{P_n}\lambda\frac{\partial P_n}{\partial p_{in}}\right]\alpha_n E_n a_{in}p_{in}^{-\sigma}P_n^{\sigma-1} \\ &= q_{in}\left(1 - \frac{p_{in} - \tau_{in}c_{in}}{p_{in}}\left[\sigma - \lambda(\sigma-1)s_{in}\right]\right), \end{aligned} \tag{1}$$

*University of Surrey, CEP and CEPR. Email: h.breinlich@surrey.ac.uk.

†Department of Economics, University of Vienna, CEPR and MaCCI. Email: harald.fadinger@univie.ac.at.

‡Department of Economics and MaCCI, University of Mannheim. Also affiliated with CEPR. Email: volker.nocke@gmail.com.

§Department of Economics and MaCCI, University of Mannheim. Also affiliated with CEPR. Email: schutz@uni-mannheim.de.

where

$$s_{in} \equiv \frac{a_{in} p_{in}^{1-\sigma}}{\sum_{j \in \mathcal{J}} a_{jn} p_{jn}^{1-\sigma}} \tag{2}$$

continues to be the market share of firm i in destination n.

Equation (1) pins down firm i's optimal markup under price competition:

$$\mu_{in} = \frac{1}{\sigma - \lambda (\sigma - 1) s_{in}},$$

where $\mu_{in} = (p_{in} - \tau_{in} c_{in})/p_{in}$ is firm i's Lerner index. Apart from this change in the expression for the firm's optimal markup, all other firm-level results go through as before.

We now turn our attention to the industry-level results. As in Section 2.3, we begin by employing an aggregative games approach to analyze the equilibrium in a given market, dropping the market subscript n to ease notation. The market-level aggregator H is now defined as

$$H \equiv P^{1-\sigma} = \sum_{j \in \mathcal{J}} a_j p_j^{1-\sigma}$$

and firm i's type as $T_i \equiv a_i (c_i \tau_i)^{1-\sigma}$.

Plugging these definitions into equation (1), making use of equation (2), and rearranging, we obtain:

$$\left(1 - s_i^{\frac{1}{\sigma-1}} \left(\frac{H}{T_i} \right)^{\frac{1}{\sigma-1}} \right) (\sigma - \lambda(\sigma - 1) s_i) = 1. \tag{3}$$

Note that the left-hand side of equation (3) is strictly decreasing in s_i on the interval $(0, \min \{\sigma/(\lambda(\sigma-1)), T_i/H\})$ and tends to σ and 0 as s_i tends to the lower and upper endpoints of that interval, respectively. Equation (3) therefore has a unique solution in the above interval, denoted $S(t_i, \lambda)$ with $t_i \equiv T_i/H$. (Solutions outside that interval necessarily give rise to strictly negative markups and are thus suboptimal.)

It is easily checked that S is strictly increasing in its first argument, strictly decreasing in its second argument, and tends to 0 and $1/\lambda$ as t_i tends to 0 and ∞, respectively.

As before, the equilibrium condition is that market shares must add up to one:

$$\sum_{j \in \mathcal{J}} S\left(\frac{T_i}{H}, \lambda \right) = 1. \tag{4}$$

The properties of the function S, described above, imply that this equation has a unique solution, $H^*(\lambda)$.

To summarize:

Proposition A. *In each destination market, and for any conduct parameter λ, there exists a unique equilibrium in prices. The equilibrium aggregator level $H^*(\lambda)$ is the unique solution to equation* (4)*. Each firm i's equilibrium market share is $s_i^*(\lambda) = S(T_i/H^*(\lambda), \lambda)$, where $S(T_i/H^*(\lambda), \lambda)$ is the unique solution to equation* (3)*.*

Proof. All that is left to do is check that first-order conditions are sufficient for optimality when $\lambda = 1$. From equation (1), we have:

$$\frac{\partial \pi_i}{\partial p_i} = q_i \left[1 - \frac{p_i - \tau_i c_i}{p_i} \left(\sigma - \lambda(\sigma - 1)\frac{a_i p_i^{1-\sigma}}{\sum_j a_j p_j^{1-\sigma}}\right)\right].$$

As the term inside square brackets is strictly decreasing in p_i, it follows that, if firm i's first-order condition holds at $\hat{p}_i$, then $\partial \pi_i / \partial p_i$ is strictly positive whenever $p_i < \hat{p}_i$ and strictly negative whenever $p_i > \hat{p}_i$. Hence, first-order conditions are sufficient for optimality. □

Having characterized the equilibrium in a given destination, we now adapt the first-order approach to industry-level gravity to the case of price competition. As in Section 2.3, let $\mathcal{E} \subsetneq \mathcal{J}$ denote the subset of exporters in country e that sell in the destination market n. The combined market share of those exporters in market n is given by

$$s_e^*(\lambda) \equiv \sum_{i \in \mathcal{E}} s_i^*(\lambda).$$

As before, we approximate $s_e^*(1)$ at the first order. The definitions of HHI and HHI_e are as in Section 2.3.

We obtain:

Proposition B. *At the first order, in the neighborhood of $\lambda = 0$, the logged joint market share in destination n of the firms from export country e is given by*

$$\log s_e^*(\lambda) = \log s_e^*(0) + \frac{\sigma - 1}{\sigma} \left[\text{HHI}(\lambda) - s_e^*(\lambda)\, \text{HHI}_e(\lambda)\right] \lambda + o(\lambda).$$

Proof. The proof follows the same developments as the proof of Proposition 2. We begin by computing the partial derivatives of S at $\lambda = 0$. It is useful to rewrite first equation (3) as

$$s_i = t_i \left(1 - \frac{1}{\sigma - \lambda(\sigma - 1)s_i}\right)^{\sigma - 1}. \tag{5}$$

Taking the logarithm and totally differentiating the equation at $\lambda = 0$ yields:

$$\frac{ds_i}{s_i} = \frac{dt_i}{t_i} - \frac{\sigma - 1}{\sigma} s_i d\lambda.$$

The partial derivatives of S are thus given by

$$t_i \partial_1 \log S(t_i, 0) = 1 \quad \text{and} \quad \partial_2 \log S(t_i, 0) = -\frac{\sigma - 1}{\sigma} S(t_i, 0).$$

To obtain $H^{*\prime}(0)$, we differentiate equation (4):

$$\sum_{j \in \mathcal{J}} \left[-\frac{T_j}{H} \partial_1 S\left(\frac{T_j}{H}, \lambda\right) \frac{dH}{H} + \partial_2 S\left(\frac{T_j}{H}, \lambda\right) d\lambda \right] = 0.$$

Setting $\lambda = 0$, plugging in the values of the partial derivatives of S, and using the fact that market shares add up to unity, we obtain:

$$\frac{H^{*\prime}(0)}{H^*(0)} = -\frac{\sigma - 1}{\sigma} \, \mathrm{HHI}(0).$$

Next, we compute $s_i^{*\prime}(0)$:

$$\begin{aligned} s_i^{*\prime}(0) &= -\frac{T_i}{H^*(0)} \partial_1 S\left(\frac{T_i}{H^*(0)}, 0\right) \frac{H^{*\prime}(0)}{H^*(0)} + \partial_2 S\left(\frac{T_i}{H^*(0)}, 0\right) \\ &= \frac{\sigma - 1}{\sigma} \left[s_i^*(0) \, \mathrm{HHI}(0) - (s_i^*(0))^2 \right]. \end{aligned}$$

Adding up and dividing by $s_e^*(0)$ yields:

$$s_e^{*\prime}(0) = \frac{\sigma - 1}{\sigma} \left[\mathrm{HHI}(0) - s_e^*(0) \, \mathrm{HHI}_e(0) \right].$$

As in the proof of Proposition 2, we can then apply Taylor's theorem to obtain the result. □

Proposition B motivates the following approximation:

$$\log s_e^*(1) \simeq \log s_e^*(0) + \frac{\sigma - 1}{\sigma} \left[\mathrm{HHI}(1) - s_e^*(1) \, \mathrm{HHI}_e(1) \right].$$

As in Section 2.3, this approximation can then be used to derive the industry-level gravity regression

$$\log \widetilde{r}_{en} = \zeta_e + \xi_n + \beta(1 - \sigma) X_{en} + \eta_{en}$$

where

$$\log \widetilde{r}_{en} \equiv \log r_{en} + \frac{\sigma - 1}{\sigma} s_{en} \, \mathrm{HHI}_{en}$$

is the value of export flows from e to n, purged from oligopolistic market power effects. Note that the correction term under price competition is equal to the one under quantity competition divided by σ. We should therefore expect to find a smaller oligopoly bias than in the main part of the paper.

C.2 Empirical Results

Table A shows results for the pooled firm-level regressions. In all specifications, the point estimates on the distance coefficient are much larger in absolute magnitude when correcting for oligopoly bias. The absolute value of the distance coefficient is slightly smaller than with Cournot competition.

Table B shows results for the pooled industry-level regressions. Again, the distance coefficient becomes larger in absolute magnitude when including the oligopoly correction term. Like in the case of Cournot competition, the absolute differences in coefficient magnitudes between the estimates with and without correction are smaller than with the firm-level estimates.

Table A: Firm-Level Gravity Estimates – Bertrand competition

	(1)	(2)	(3)	(4)
Method	PPML w/o corr	PPML w/ corr $\sigma = 4$	PPML w/corr $\sigma = 5$	PPML w/ corr $\sigma = 6$
log dist.	-0.874***	-1.307***	-1.418***	-1.473***
	(0.021)	(0.156)	(0.190)	(0.207)
Obs.	11,955,786	11,955,786	11,955,786	11,955,786
(Pseudo) R^2	0.14	0.23	0.26	0.27
Firm-ind.-year FE	YES	YES	YES	YES
Ind.-dest.-year FE	YES	YES	YES	YES

Notes: Firm-level data, pooled across industries and years. Results for top 3 exporters. Oligopoly correction with $\sigma \in \{4, 5, 6\}$. Standard errors in parentheses, clustered at the destination-year level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B: Industry-level Gravity Estimates – Bertrand Competition

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Meth.	Heck	Heck	Heck	Heck	HMR^2	HMR^2	HMR^2	HMR^2
	w/o	w/	w/	w/	w/o	w/	w/	w/
		$\sigma = 4$	$\sigma = 5$	$\sigma = 6$		$\sigma = 4$	$\sigma = 5$	$\sigma = 6$
log dist.	-1.189***	-1.215***	-1.217***	-1.218***	-1.151***	-1.176***	-1.177***	-1.178***
	(0.198)	(0.202)	(0.202)	(0.202)	(0.190)	(0.194)	(0.194)	(0.194)
inv. Mills	-0.121	-0.123	-0.123	-0.123	0.678***	0.700***	0.702***	0.703***
	(0.165)	(0.168)	(0.168)	(0.168)	(0.169)	(0.173)	(0.174)	(0.174)
$\log \hat{Z}$					0.907***	0.937***	0.939***	0.940***
					(0.265)	(0.270)	(0.270)	(0.270)
$\log \hat{Z}^2$					-0.103*	-0.107*	-0.108*	-0.108*
					(0.057)	(0.058)	(0.058)	(0.058)
Obs.	60,662	60,662	60,662	60,662	60,662	60,662	60,662	60,662
R^2	0.302	0.297	0.2.97	0.297	0.304	0.299	0.299	0.299
Ind-orig FE	YES	YES	YES	YES	YES	YES	YES	YES
Ind-dest FE	YES	YES	YES	YES	YES	YES	YES	YES

Notes: Industry-level data, pooled across industries, second step of HMR procedure. Standard errors clustered at destination level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

D Supplementary Tables

Table C: Firm-Level Gravity Estimates. OLS

	(1)	(2)	(3)	(4)
Method	OLS w/o corr	OLS w/ corr	OLS w/corr	OLS w/ corr
		$\sigma = 4$	$\sigma = 5$	$\sigma = 6$
log dist.	-0.232***	-0.264***	-0.275***	-0.285***
	(0.014)	(0.015)	(0.015)	(0.015)
Obs.	708,392	708,392	708,392	708,392
R^2	0.06	0.05	0.05	0.04
Firm-ind.-year FE	YES	YES	YES	YES
Ind.-dest.-year FE	YES	YES	YES	YES

Notes: Firm-level data, pooled across industries and years. Results for top-3 exporters, without and with oligopoly correction. Standard errors in parentheses, clustered at the destination-year level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table D: Firm-Level Gravity Estimates. PPML Without Zeroes

	(1)	(2)	(3)	(4)
Method	PPML w/o corr	PPML w/ corr	PPML w/corr	PPML w/ corr
		$\sigma = 4$	$\sigma = 5$	$\sigma = 6$
log dist.	-0.410***	-1.219***	-1.261***	-1.279***
	(0.017)	(0.334)	(0.352)	(0.360)
Obs.	708,392	708,392	708,392	708,392
(Pseudo) R^2	0.09	0.32	0.33	0.33
Firm-ind.-year FE	YES	YES	YES	YES
Ind.-dest.-year FE	YES	YES	YES	YES

Notes: Firm-level data, pooled across industries and years, and dropping observations with zero trade flows. Results for top-3 exporters, without and with oligopoly correction. Standard errors in parentheses, clustered at the destination-year level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table E: Firm-level Gravity Estimates. Robustness on Sample of Firms

	(1)	(2)	(3)	(4)	(5)	(6)
Sample	Top 1	Top 1	Top 3	Top 3	Top 5	Top 5
Method	PPML w/o	PPML w/ $\sigma = 5$	PPML w/o	PPML w/ $\sigma = 5$	PPML w/o	PPML w/ $\sigma = 5$
log dist.	-0.978***	-1.257***	-0.874***	-1.518***	-0.793***	-1.532***
	(0.018)	(0.083)	(0.021)	(0.220)	(0.021)	(0.232)
Obs.	3,690,099	3,690,099	11,955,786	11,955,786	20,265,693	20,265,693
(Pseudo) R^2	0.14	0.25	0.14	0.28	0.14	0.29
Firm-ind.-year FE	YES	YES	YES	YES	YES	YES
Dest.-ind.-year FE	YES	YES	YES	YES	YES	YES

Notes: Firm-level data, pooled across industries. Results for top 1 exporters (columns 1–2), top 3 exporters (columns 3–4), top 5 exporters (columns 5–6). Standard errors in parentheses, clustered at the destination-year level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table F: Monte Carlo: Firm-Level Gravity Estimates. OLS and PPML Without Zeroes

	(1)	(2)	(3)	(4)
Method	OLS w/o corr	OLS w/ corr $\sigma = 5$	PPML w/o corr	PPML w/ corr $\sigma = 5$
log dist.	-0.454**	-0.472**	-0.717***	-1.024***
	(0.202)	(0.214)	(0.201)	(0.362)
Obs.	9,529	9,529	9,529	9,529
(Pseudo) R^2	0.48	0.47	0.21	0.65
Firm-run FE	YES	YES	YES	YES
Dest.-run FE	YES	YES	YES	YES

Notes: Monte Carlo dataset, pooled across Monte Carlo runs. Results for top-3 exporters, without and with oligopoly correction. PPML results when dropping observations with zero trade flows. Standard errors in parentheses, clustered at the destination level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table G: Monte Carlo: Firm-level Results. Robustness on Sample of Firms

	(1)	(2)	(3)	(4)	(5)	(6)
Sample	Top 1	Top 1	Top 3	Top 3	Top 5	Top 5
Method	PPML w/o	PPML w/ $\sigma = 5$	PPML w/o	PPML w/ $\sigma = 5$	PPML w/o	PPML w/ $\sigma = 5$
log dist.	-1.093***	-2.480***	-0.851***	-1.403***	-0.805***	-1.421***
	(0.267)	(0.703)	(0.135)	(0.365)	(0.104)	(0.310)
Obs.	6,340	6,340	30,198	30,198	55,680	55,680
(Pseudo) R^2	0.28	0.80	0.21	0.74	0.19	0.71
Firm-run FE	YES	YES	YES	YES	YES	YES
Dest.-run FE	YES	YES	YES	YES	YES	YES

Notes: Monte Carlo dataset, firm-level data, pooled across Monte Carlo runs. Results for top 1 exporters (columns 1–2), top 3 exporters (columns 3–4), top 5 exporters (columns 5–6). Standard errors in parentheses, clustered at the destination level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.