

Hirte, Georg; Lessmann, Christian

Working Paper

Trade, Integration, and Interregional Inequality

CESifo Working Paper, No. 4799

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Hirte, Georg; Lessmann, Christian (2014) : Trade, Integration, and Interregional Inequality, CESifo Working Paper, No. 4799, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/102259>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



Working Papers

www.cesifo.org/wp

Trade, Integration, and Interregional Inequality

Georg Hirte
Christian Lessmann

CESIFO WORKING PAPER NO. 4799
CATEGORY 8: TRADE POLICY
MAY 2014

An electronic version of the paper may be downloaded

- *from the SSRN website:* www.SSRN.com
- *from the RePEc website:* www.RePEc.org
- *from the CESifo website:* www.CESifo-group.org/wp

Trade, Integration, and Interregional Inequality

Abstract

We study the effect of international trade and freeness of trade (openness) on interregional inequality within countries. We estimate a model derived from a structural economic-geography approach in which interregional inequality depends on weighted trade shares and trade costs. In addition to the standard trade-to-GDP ratio, we derive and propose an aggregate freeness-of-trade measure based on phiness of trade. Both measures are instrumented by proxies constructed from estimates of a gravity model of bilateral trade, which covers 208 countries for the period 1948-2006. For our study we use Gennaioli et al.'s (2013) cross-country data set, which covers 110 countries (1569 sub-national regions) for the year 2005, and the panel data set of Lessmann (2014), which covers 56 countries (835 sub-national regions) for the period 1980-2009. The IV and dynamic panel regressions provide evidence that trade increases interregional inequality, but that the coefficient of the freeness-of-trade variable is ambiguous. Because the latter is an indicator for integration in the world markets, we conclude that more integration may neutralize the negative interregional distribution effects of trade.

JEL-Code: R120, F100, C360.

Keywords: regional inequality, trade, gravity model, New Economic Geography, panel data.

Georg Hirte
Technical University Dresden
Institute of Transport & Economics
Chair of Spatial Economics &
Regional Science
Germany – 01062 Dresden
georg.hirte@tu-dresden.de

Christian Lessmann
Technical University Carolo-Wilhelmina
Braunschweig
Institute of Economics
Germany – 38106 Braunschweig
c.lessmann@tu-braunschweig.de

This version: April 08, 2014

We are grateful to Andrei Shleifer, Rafael La Porta, and their coauthors for sharing their regional data with us. Moreover, we would like to thank the participants in the Winter Seminar of the German-speaking section of ERSa in Matrei, the International Economic Policy Research Seminar at the Goethe University Frankfurt, the ERSa conference in Palermo, and the regional economists' meeting of the German Economic Association for many helpful comments.

1 Introduction

For a long time economists have been working on the effect of trade on inequality within countries. Most of this research is on intra-country wage inequality (overview in Pavcnik, 2013), but since the advent of the New Economic Geography, spatial effects (review by Brühlhart, 2011), intra-group inequality (e.g., Egger and Kreickemeier, 2012), and interregional inequality (e.g., Brühlhart et al., 2012) have aroused the interest of scholars. We pick up this latter issue and study whether international openness contributes to interregional income inequality within countries.

A first glance at data reveals that openness and inequality are correlated in a number of countries. Take India and Bolivia as examples. Whereas the per capita GDP of the poorest federal state in India, Bihar, was 24% of that of the richest state, Delhi, in 1980, its relative income decreased to 13% in 2005. In Bolivia, our second example, the income of the poorest department, Potosí, was 43% of that of the richest department, Santa Cruz, in 1988, but the relative performance of Potosí decreased to 29% in 2009. A possible cause of this development is international trade. Both countries have significantly increased their integration into international trade. In India, the ratio of the sum of exports and imports to GDP, i.e., the trade-to-GDP ratio, increased from 15.6% in 1980 to 41.2% in 2005, whereas in Bolivia it increased from 41.9% in 1988 to 64% in 2009. This positive correlation between the trade-to-GDP ratio and interregional inequality gives a first hint at a possible link between the two variables.

A high level of regional inequality is of major concern because it may cause serious economic and political trouble. Recent studies show that interregional inequality is a breeding ground for separatist movements and a major determinant of internal conflicts, such as civil war (e.g. Østby, 2009, Buhaug et al., 2012, Lessmann, 2013). Further, evidence suggests that interregional inequality determines about one third of interpersonal inequality (Yemtsov, 2005, Elbers et al., 2005). In light of these findings, it is important to understand the determinants of regional inequality.

In this paper we ask whether international trade and openness to trade are among these determinants. We examine this issue for a large number of countries all over the world and derive our empirical model from a structural economic-geography approach.

The theoretical literature is ambiguous concerning the direction of the effects (see the overview by Brühlhart, 2011). In the traditional trade theory, trade increases regional inequality if the initially poorer regions of a country gain relatively less from the opening to the world market and

the rich regions relatively more (Wood, 2002, Anderson, 2005)¹. In models of the New Economic Geography the effect depends on other conditioning factors such as trade costs (Krugman and Livas Elizondo, 1996, Paluzie, 2001, Behrens et al., 2006), the initial level of inequality (Alonso Villar, 1999), the initial country size (Zhang and Zhang, 2010), or the strength of the market crowding effect (Brühlhart et al., 2004, Crozet and Koenig-Soubeyran, 2004).

Despite the fact that theory does not provide unambiguous signs, most empirical studies focusing on single countries provide evidence in favor of a positive link between trade and interregional inequality (e.g., Chiquiar, 2008, on Mexico, Zhang and Zhang, 2003, on China, Brühlhart et al., 2013, on Austria).² But one cannot infer from these studies whether this outcome is in general a result of trade.

There are also some studies on panels of countries. For instance, Egger et al. (2005) found evidence that trade raises interregional wage inequality in a panel of eight central and eastern European countries (CEECs) over 1991–1999. More specifically, they found evidence that intermediate-good exports significantly contribute to the increase in the standard deviation of regional wages. In contrast, Milanovic (2005) finds no significant coefficient for the five most populated countries in the period 1978–2000 in a similar study. Petrakos et al. (2005) examined the effect of growth and integration on interregional inequality in eight European countries from 1981 to 1997. They found that integration, measured as the ratio of intra-EU trade to GDP, raises inequality in France and Spain, lowers it in the Netherlands and Portugal, and is insignificant for Greece, Italy, and the UK.

Barrios and Strobl (2009) examined a panel of the EU 15 countries for 1970–2000. They found evidence that the coefficient of the real openness measure is positively significant, suggesting that real trade openness raises interregional inequality in the EU15. More recently, Rodríguez-Pose (2012) studied an unbalanced panel of 28 countries for 1980–2005 and provided evidence that only interaction terms of trade with development and “coincidence” measures are statistically significant, while the coefficient of trade is insignificant.³ This literature uses a rather limited

¹Other channels are migration (Haaparanta, 1998; evidence by Chiquiar, 2008), the distribution of factor ownership (e.g. Anderson, 2005), and the willingness to redistribute via taxes and transfers (Rodrik and van Ypersele, 2001).

²See also Chiquiar (2008), González Rivas (2007), Hanson (1997, 1998), and Sanchez-Reaza and Rodríguez-Pose (2002, 2005) on Mexico; Kanbur and Zhang (2005) on China; Pernia and Quising (2003) on the Philippines; Breau and Rigby (2010) on Canada; and Volpe (2010) on Brazil.

³There is another study, by Rodríguez-Pose and Gill (2006), using an unbalanced panel for eight countries in 1970–2000. They, however, only use a graphical representation of the trade content index.

number and differing selection of countries implying a high variability of the estimation results.

In a recent paper, Ezcurra and Rodríguez-Pose (2013) examine the effect of integration measured by the KOF index of Globalization on regional inequality in a panel of 47 countries over the period 1990–2007. They apply pooled OLS estimates and find a positive coefficient for globalization to be higher in medium- and low-income countries. However, they do not take account of endogeneity issues and of unobserved heterogeneity of countries, and their database comprises only a very small number of LDCs.

This literature has left several open issues: (i) The size of the cross-country database is usually very small, biased in favor of industrial countries and hardly considering developing countries. As a consequence, a selection bias may occur. (ii) With a few exceptions (e.g., Barrios and Strobl, 2009, Brühlhart et al., 2013), the empirical studies are not based on sound theoretical models. (iii) Most empirical studies use the trade-to-GDP ratio as a proxy for openness and thus do not consider important problems concerning openness and integration that are present in the theory.

We deal with these issues in the following:

(i) There are two recently published cross-country databases for us to take advantage of. Genaioli et al. (2013) provide a cross-section sample of regional income data for 110 countries with 1569 sub-national regions for the year 2005, while Lessmann’s (2014) panel database comprises 56 countries with 835 sub-national regions for the period 1980–2009. Therefore, we can consider a considerably larger number of countries than in any study on that topic before, and hence reduce the selection bias present in other studies and provide much more general evidence. In our theoretical model trade is endogenously determined by the spatial structure of firms and populations, depending on differences in the real return to entrepreneurs. Because regional consumer price indices are not observable, there is an omitted-variable bias. In addition there is a causality bias, since differences in the GDP per capita affect the differences in the real reward and as a consequence influences the relocation of firms. Further, we do not have information on interregional trade costs. If we add unobserved heterogeneity (e.g., amenities) to the model, this provides another reason why the covariance between trade (or phiness; see below) and the error term deviates from zero.

Unfortunately, the database of Genaioli et al. (2013) includes only observations in 2005 and is therefore not suitable for panel data analysis. To be able to exploit this unique database we carefully address the endogeneity problem and use two strategies to cope with it. First, we construct an instrument for the real trade-to-GDP ratios from an estimated gravity model of bilateral trade

(Frankel and Romer, 1999, and Alcalá and Ciccone, 2004). Then, we combine the bilateral-trade data for more than 200 countries of Head et al. (2010) with the estimation approaches suggested by Frankel and Romer (1999), Anderson and van Wincoop (2003), and Baier and Bergstrand (2007). Thereby, we augment the gravity equation with some standard determinants of trade, following Head et al. (2010), and add country-pair fixed effects in order to deal with multilateral resistances (Anderson and van Wincoop, 2003). Then we use the constructed trade shares in the IV estimates as instrument for the trade-to-GDP ratio. As our second approach, we employ dynamic panel regressions using a system generalized method of moments (GMM) estimator as in Arellano and Bover (1995) and Blundell and Bond (1998) to the Lessmann (2014) database.

(ii) Our regression models including the gravity approaches derive from a structural New Economic Geography model that is an extension of the approach of Pflüger (2004). In that way, we reconcile the empirical approaches with the respective theories, unlike previous cross-country studies in that context.

(iii) While the literature predominantly uses the trade-to-GDP ratio, we propose to apply another measure for openness. The trade-to-GDP ratio is commonly used as a proxy for the external-market potential, though it does not explicitly consider the freeness of trade, which is one of the central ingredients of that potential. Consequently, we propose and use an aggregate freeness-of-trade index based on bilateral openness as calculated according to Head and Ries (2001)⁴. The freeness-of-trade measure focuses on integration in the world market and thus implicitly considers regional free-trade agreements, etc. Since trade and openness might be endogenous, as will free-trade agreements, membership in the OECD, etc., we instrument the openness measure too.⁵

We find evidence that trade raises interregional inequality. The point estimate in the cross-country regressions implies that an increase in the trade-to-GDP ratio by 10 percentage points is on average associated with an increase in regional inequality by approximately 2%. A variety of robustness checks, including semi-parametric estimates, confirm our findings. In contrast, we cannot identify a significant coefficient of freeness of trade in the panel regressions, though it is significant in the IV cross-section estimates. This hints at the relevance of integration. If we interpret our measure of freeness of trade as a measure of integration, we can conclude that the higher integration is in the world markets, the less harmful is more trade to interregional inequality.

⁴By using this economic measure of trade openness we implicitly also consider de jure trade openness (e.g., Sachs and Warner, 1995).

⁵Arrabias et al. (2009) provide and discuss further measures.

The reminder of the paper is organized as follows: Section 2 presents the simple New Economic Geography model and derives our regression equation. Section 3 presents our econometric analysis. Thereafter, we first discuss our measure of regional inequality (section 3.1.1), present our empirical approach, and develop our instrument of trade shares (section 3.3). Subsequently we present our main results from cross-section and panel regressions. Then we provide different robustness tests, including alternative measures of regional inequality as dependent variable, semi-parametric regression results, and different interaction variables. Section 4 summarizes our findings and concludes.

2 Theory

2.1 A simple model

Since interregional inequality might depend on the spatial allocation of industry and households, a model that takes account of this link is required. Therefore, we derive the empirical model as well as the gravity approach from a model in the tradition of the New Economic Geography (Krugman, 1991). In particular, we adopt the two-region model of Pflüger (2004) and adjust it to our purpose. We use this model because it is able to handle all kinds of spatial allocation of industry and population and thus is better suited than models featuring only full dispersion or full agglomeration. Nevertheless, it is easy to use, though it preserves almost all features of the standard core-periphery model (Krugman, 1991).

There are two countries i and j , where country i is called Home. We focus on Home, which consists of S regions with indices r or s . Though we assume full symmetry between the two countries, we simplify equations by considering the other country, called Foreign, as a black box. It is straightforward to extend the model to take account of the internal structure of Foreign. There are two types of goods: (i) a homogeneous tradable good, Z , produced with constant returns to scale, a unit demand of the sole input labor and zero trade costs, and (ii) a mass of varieties of a monopolistic good produced with one unit of skilled labor per variety as a fixed input and c units of labor per output as a variable input. We normalize the wage to unity; then, due to the construction of the homogeneous-good sector, the price of these goods is also unity. In the following we focus on Home and simplify notation by dropping the index i when we look at country i 's regions indexed

r or s . Aggregate variables of Home or Foreign are indexed by i .

Home is endowed with a fixed supply of skilled labor or entrepreneurs, K , and each region with a fixed amount of other types of labor, L_r , simply called labor. Regional skilled labor is endogenous on account of migration. The regional population is given by $P_r = L_r + K_r$.

Each household in a region consumes the local good Z_r and varieties m of the aggregate monopolistic good M_r . Upper-level utility is represented by the quasi-linear utility function (Pflüger, 2004) and sub-utility by a Spence–Dixit–Stiglitz utility function with love for variety; thus,

$$U_r = Z_r + \alpha_r \ln M_r, \quad M_r = \left(n_r m_{rr}^{\frac{\sigma-1}{\sigma}} + n_s m_{sr}^{\frac{\sigma-1}{\sigma}} + n_j m_{jr}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \quad (1)$$

where α denotes expenditures on M , m_{sr} is the demand for a variety produced in region s , and σ is the elasticity of substitution, which also represents the price elasticity of demand. n_r is the mass of varieties produced in region r , n_i the mass produced in Home, and a mass $n_j = N - n_i$ is produced abroad. The budget constraint is $G_r = Z_r + Q_r M_r$, where G_r denotes household income and Q_r is the perfect price index for monopolistic goods.

Transport of monopolistic goods is subject to iceberg transportation costs. Following Behrens et al. (2007), transportation costs refer to intra-country transportation costs, and trade costs to international costs. Transportation costs between regions, Ψ , are assumed to be symmetric, and $\Psi > 1$ per unit of a variety. We assume that trade costs are symmetric for imports and exports, and that there is no distinction concerning the region of destination or origin in Foreign, but that they might be different among regions of country i . Let $T_r > 1$ represent the trade costs of Home's region r to Foreign.

The demand functions for the varieties are

$$m_{rr} = \alpha_r \frac{Q_r^{\sigma-1}}{q^\sigma}, \quad m_{sr} = \alpha_r \frac{Q_r^{\sigma-1}}{\Psi^\sigma q^\sigma}, \quad m_{jr} = \alpha_r \frac{Q_r^{\sigma-1}}{T_r^\sigma q^\sigma}. \quad (2)$$

In the monopolistic sector a mass of goods is produced by identical firms with an increasing-returns-to-scale technology giving rise to Chamberlin monopolistic competition. The mass of firms, n_r , is equal to the mass of entrepreneurs, K_r , we can use n instead of K below, and λ_r determines the share of region r in all firms of country i . Since firms are identical, we drop the indices of firms and regions. Further, $\lambda_r = n_r/n_i$ and $\mu_i = n_i/N$ denote regional shares in the national number

of firms and country i 's share on the global number of firms, respectively. Because we consider a mass of firms, profit maximization implies that the mill price is a mark-up on marginal costs, i.e.,

$$q = \frac{\sigma - 1}{\sigma} c. \quad (3)$$

Accordingly, the perfect price indices are

$$Q_r = Nq [\mu_i (\lambda_r + \psi \lambda_s) + \phi_r \mu_j]^{\frac{1}{1-\sigma}}, \quad Q_j = Nq \left(\mu_i \sum_s \phi_s \lambda_s + \mu_j \right)^{\frac{1}{1-\sigma}}, \quad (4)$$

where $\psi \equiv \Psi^{1-\sigma} \in [0, 1]$ is the index of internal integration while $\phi_r \equiv T_r^{1-\sigma} \in [0, 1]$ represents trade freeness between Home's region r and Foreign.

Because operating profits are zero, R equals the fixed costs, i.e., $R = (q - c)x$. Hence, a firm's scale is $x = (\sigma - 1)(R/c)$ and its cost function is $\sigma R = R + cx$, implying $\sigma R = qx$. Clearing of the market of a variety implies that revenue equals the value of sales to all regions, i.e.,

$$\sigma R = qx_r = P_r q m_{rr} + q P_s \Psi m_{rs} + P_j T_r m_{rj}. \quad (5)$$

After substituting (2) and (4) into (5), we see that the factor income in manufacturing is a function of the nominal market access (MA) of the region. The latter is the sum over local market access, interregional market access, and international market access; thus⁶

$$\sigma R_r = \underbrace{P_r \alpha_r \left(\frac{Q_r}{q} \right)^{\sigma-1}}_{\text{local MA}} + \underbrace{\psi P_s \alpha_s \left(\frac{Q_s}{q} \right)^{\sigma-1}}_{\text{interregional MA}} + \underbrace{\phi_r P_j \alpha_j \left(\frac{Q_j}{q} \right)^{\sigma-1}}_{\text{international MA}}. \quad (6)$$

This is the region's income from manufacturing. Full regional income requires us to add labor employed in production of homogeneous tradables⁷. In the following we derive differences in per capita income for manufacturing⁸.

In addition there is some kind of dynamics in the model because λ is endogenous. Entrepreneurs

⁶The core-periphery model of Krugman (1991), like other NEG approaches, produces the same general pattern of this equation. Therefore, the following holds for all that kind of models.

⁷Redding and Venables (2004) solve for the marginal costs, which are included in q above. These marginal costs (factor incomes) depend on the real market access divided by σ . Our solution refers to zero profits. Thus σR represents rewards from skilled and unskilled labor in manufacturing. Then $(\sigma - 1)R$ is equivalent to wages paid to the variable factor. For this reason, it is not necessary to solve for q , and the RHS represents the real market access.

⁸Extending the approach to include income from homogeneous tradables is straightforward.

move into the region offering them the higher real reward, i.e., the higher indirect utility. Therefore λ depends on all parameters of the model.

2.2 Income inequality

Equation (6) is the basis for the whole econometric analysis. We use this equation in several ways. First, we rewrite the equation to determine inequality according to different definitions of trade openness.

The standard approach to get a link between market access and trade openness is to assume that the trade-to-GDP ratio is a proxy for international market access divided by Ricardian profits (e.g., Redding and Venables, 2004). This, however, is not fully consistent with the a typical NEG model such as the one we use. Market access in (6) refers to exports. To consider imports we have to extend this equation (see below). Second, freeness of trade, or phiness (ϕ) is an important determinant of regional income. It is thus much more natural to use phiness as measure of openness in these models.

Rearranging (6) will show how freeness of trade determines interregional inequality. Another problem arises when we switch to a world with more than two countries. Then bilateral openness has to be extended to a multilateral openness measure. The following exercise will provide a multilateral phiness index. Later on we can also derive a gravity equation to estimate and construct the instruments for both openness measures.

2.2.1 Inequality measures using the trade-to-GDP ratio

For the time being we assume that workers are equally distributed across regions and that technologies are identical. Since we further assume constant marginal productivity of labor, income differences across regions depend on Ricardian profits. We denote the per capita (p.c.) income of region r by $y_r = (n_r \sigma R_r) / P_r$ and define y_i as Home's average income p.c. Next, to simplify the notation we denote the local market access by A_{rr} , the market access of r in s by A_{rs} , and the international market access (i.e., exports to all countries) by E_r . These are

$$A_{rr} \equiv n_r P_r \alpha_r \left(\frac{Q_r}{q} \right)^{\sigma-1}, \quad A_{rs} \equiv \Psi_{rs} n_r P_s \alpha_s \left(\frac{Q_s}{q} \right)^{\sigma-1}, \quad E_r \equiv n_r T_{rj} P_j \alpha_j \left(\frac{Q_j}{q} \right)^{\sigma-1}. \quad (7)$$

We use these definitions to rewrite the region's expenditure for varieties as well as (6), thus obtaining the two accounting equations for the use of income and the use of production:⁹

$$Y_r = A_{rr} + A_{sr} + I_r, \quad Y_r = A_{rr} + A_{rs} + E_r.$$

Adding these, aggregating over all regions dividing by P_i yields the country's per capita income depending on overall trade

$$y_i = \frac{A_{rr} + A_{ss}}{P_i} + \frac{A_{rs} + A_{sr}}{P_i} + \frac{1}{2} \frac{I_i + E_i}{P_i}. \quad (8)$$

A region's relative deviation in per capita income from the average is then

$$\frac{y_r - y_i}{y_i} = \left(\frac{e_{Ar}}{p_r} - 1 \right) + \frac{1}{2p_r} [(e_{Ir} - p_r e_{Ii}) + (e_{Er} + p_r e_{Ei} - 2e_{Ar} e_{Er} + p_r)] T_i, \quad (9)$$

where I represents imports and E exports. Here e_{Ir} and e_{Er} are region r 's import and export shares in the country's aggregate trade $I_i + E_i$, respectively, e_{Ii} and e_{Ei} are the shares of imports and exports in the trade of country i , and T_i is the trade-to-GDP ratio of the country. Further, we have used the definitions for absorption, the region's share in absorption, and the link between exports and absorption given by

$$A_i \equiv A_{rr} + A_{ss} + A_{rs} + A_{sr}, \quad e_{Ar} \equiv \frac{A_{rr} + \frac{1}{2}A_{rs} + \frac{1}{2}A_{sr}}{A_i}, \quad \frac{A_i}{Y_i} = \frac{Y_i - E_i}{Y_i}.$$

Equation (9) states that a region's relative income depends on two components. The first term is a measure of intra-country openness. If a region's share of all intra-country trade, divided by its population share, is larger than that of the average region, the region is relatively rich. In this case it trades more within itself or with other regions, implying gains from intra-country trade. In addition the markup on prices, and thus the income, depends on the tightness of competition in that region. The second term is the weighted trade-to-GDP ratio of the country, where the weights depend on the deviation of regions' shares of imports and exports from the country shares.

⁹As we have shown above, we can use nominal export values, due to the link between Ricardian profits and factor income. This is in contrast to the approach used by Egger et al. (2005), Redding and Venables (2004), or Head and Mayer (2006), who use real exports.

This implies the econometric model

$$I_{it} = b_{1k}\mathbf{Z}_{ikt} + b_{2h}\mathbf{X}_{iht} + b_3T_{it} + b_{4h}\mathbf{X}_{iht} \times T_{it} + a_i + d_t + \varepsilon_{it}, \quad (10)$$

where I_{it} is a measure for interregional inequality in country i at time t , \mathbf{Z}_{ikt} are exogenous control variables, T_{it} is the trade-to-GDP ratio, \mathbf{X}_{ikt} are other control variables that interact with this measure and represent the weights in (9), d_t are time effects, a_i are country fixed effects, and ε_{it} is the error term. In the specifications that use the cross-section data set, we estimate a similar model, which, of course, has no fixed effects and no time dimension.

2.2.2 Inequality measure using the freeness of trade

In our model, freeness of trade enters all parts of the inequality equation (9) and thus is an important determinant of interregional inequality, despite the fact that phiness (Baldwin et al., 2005) it is not used in the empirical studies on interregional inequality.¹⁰

We substitute the price indices (6) into the equation for operating profits (6) and divide by P_r to get

$$y_r = N^\sigma \mu_i \frac{\lambda_r}{p_r} \left(\frac{p_r \alpha}{\Delta_r} + \psi \frac{p_s \alpha}{\Delta_s} + \phi_{rj} \frac{P_i \alpha_j}{P_i \Delta_j} \right), \quad (11)$$

where

$$\Delta_r \equiv \mu_i (\lambda_r + \psi \lambda_s) + \phi_{rj} \mu_j, \quad \Delta_j \equiv \mu_i \sum_s \phi_s \lambda_s + \mu_j$$

are terms inversely related to the price indices. A country's income per capita is given by aggregating (6) and dividing by P_i (see Appendix):

$$y_i = N^\sigma \mu_i \sum_s \lambda_s \left(\frac{p_s \alpha}{\Delta_s} + \psi \frac{p_k \alpha}{\Delta_k} + \phi_{sj} \frac{P_j \alpha_j}{P_i \Delta_j} \right). \quad (12)$$

Manipulating the difference eventually yields (see Appendix)

$$\frac{y_r - y_i}{y_i} = \left(\frac{e_{Er}}{p_r} - \frac{e_{Ar}}{p_r} \right) \frac{E_i}{Y_i} = N^\sigma \left[\frac{\lambda_r}{p_r} - \frac{\frac{\lambda_r p_r}{\Delta_r} + \psi \frac{\lambda_s p_s}{\Delta_s p_r}}{(1 + \psi) \sum_s \frac{\lambda_s p_s}{\Delta_s}} \right] \Phi_i, \quad (13)$$

where γ_i is the ratio of Home's GDP to the aggregate GDP of all destination countries. Further,

¹⁰Combes et al. (2005) provide an example of the empirical use of bilateral freeness of trade in a study on another topic. Bosker and Garretsen (2010) emphasize that the way trade costs are modeled matters in empirical studies. They also propose to consider freeness of trade.

we assume that the ratio of region r 's trade costs to the country's trade cost is constant across all destination countries. We now have a measure for the aggregate freeness of trade of Home:

$$\Phi_i = \sum_j \frac{\gamma_j}{\gamma_i} \frac{\mu_i \phi_{ij}}{\mu_i \phi_{ij} + \mu_j} \equiv \sum_j \frac{y_j}{y_i} \frac{\phi_{ij}}{1 + \frac{P_i}{P_j} \phi_{ij}}, \quad (14)$$

where γ_j is the ratio of country j 's GDP to the aggregate GDP of all destination countries, and ϕ_{ij} are symmetric bilateral trade costs between countries i and j . Interregional differences in the income depend now on the weighted aggregate freeness of trade where the weights refers to the regional freeness of trade, the region's population share, and the region's production's share in the country's absorption.

After linearizing, we get the econometric model

$$I_{it} = a_i + b_{1kt} Z_{ikt} + b_2 \Phi_{it} + d_t \epsilon_i, \quad (15)$$

where $a = N^\sigma$, Z_{ik} is a vector of controls for the factor in square brackets in (13), Φ_i is the overall freeness of trade of country i defined in (14), and I_i is the logarithm of the income distribution measure. Unfortunately, bilateral trade openness also enters the price indices (13) and determines λ . For this reason we have to test linearity and also consider interaction terms below.

3 Empirical Analysis

3.1 Data and Variables

3.1.1 Interregional inequality

Below we measure regional inequality (I) by the population-weighted coefficient of variation of the regional GDP per capita as calculated by

$$I_{it} := \frac{1}{\bar{y}_{it}} \left[\sum_{r=1}^S p_{rit} (y_{rit} - \bar{y}_{it})^2 \right]^{1/2}, \quad (16)$$

where \bar{y}_{it} is country i 's average GDP p.c. in period t , y_{rit} is the per capita income of region r in country i , p_{rit} is the share of the country's total population in region r , and n is the number of spatial units. This measure has frequently been used in economic geography, and was introduced

by Williamson (1965) (see Lessmann, 2013, for detailed descriptions of the data). The measure is mean-independent, is independent of the sizes and the number of spatial units, and satisfies the Pigou–Dalton transfer principle, which states that a transfer from rich to poor regions should reduce the inequality measure (see Dalton, 1920, Pigou, 1912, and Sen, 1973, for details).

We use two different data sets to calculate the measures of regional inequality. First, we refer to Lessmann (2014), who has collected a panel data set of regional inequality. It covers 56 countries (835 sub-national regions) for the period 1980–2009. Second, we use the regional data provided by Gennaioli et al. (2013). This cross-section data set covers 110 countries (1569 sub-national regions) for the year 2005. Both data sources have their strengths and weaknesses. The data provided by Gennaioli et al. (2013) covers more countries, in particular less-developed countries, which brings more variation to the cross section. But this data set has no panel structure, which Lessmann (2014) does. Therefore, we decided to use both data sets and to compare the results.

The territorial level at which the regional income is measured differs slightly between Lessmann (2014) and Gennaioli et al. (2013). While Lessmann (2014) concentrates on NUTS2 regions for European countries and on states and provinces otherwise, Gennaioli et al. (2013) concentrates on the sub-national level with the highest political authority. Therefore, the regions are more heterogeneous in size in the data set of Gennaioli et al. (2013). Another small difference is the measure of regional income: while Lessmann (2014) uses the regional per capita GDP, Gennaioli et al. (2013) also include regional wage data and data on household income where GDP data is not available. The correlation between the inequality measures based on the different data sources is fairly high (0.7724).

To get a first impression of the data, **Table 1** shows our inequality measure for all countries considered in the cross section. We have grouped the countries by their gross national income per capita using the 2013 World Bank classification. Most countries in our data set are high-income countries, but we have also a good representation of low- and middle-income countries. Interestingly, high-income countries have on average lower regional inequality than upper-middle-income countries, and upper-middle-income countries have higher regional inequality than lower-middle-income countries. This is in line with the theory of Kuznets (1955) and Williamson (1965), which suggests an inverted U-shaped relationship between regional inequality and development (see Barrios and Strobl, 2009, and Lessmann, 2014, for empirical evidence for the Kuznets curve in regional inequality). In contrast with these findings, a smaller number of low-income countries have

higher regional inequality. However, if we merge the groups of low-income countries and lower-middle-income countries, we obtain an average regional inequality of 0.38, which again supports the inverted-U hypothesis.

Note that the quality of the regional data in low-income countries is not very high. Gennaioli et al. (2013) have to make several adjustments to the data to make it compatible with country-level income data as reported by the World Development Indicators; therefore, we should be cautious when interpreting the data on low-income countries.

3.1.2 Trade and freeness-of-trade measures

T_i is the trade-to-GDP ratio (e.g., Redding and Venables, 2004). We measure it as the ratio of the sum of exports and imports to the GDP of country i in purchasing power parities (see Alcalá and Ciccone, 2004). The actual trade shares come from the World Development Indicators Series, while the trade-to-GDP ratio is provided by Penn World Tables 7.1 (Heston et al., 2012).

To handle endogeneity we apply IV estimates. We first estimate the trade-to-GDP ratio, \hat{T}_i , by applying a gravity approach to bilateral trade. The trade instrument is calculated using the data by Head et al. (2010), who refer to the International Monetary Fund's Direction of Trade Statistics (DOTS). Subsequently, we apply an IV estimation with \hat{T}_i as instruments for the trade-to-GDP ratio.

When we use the second approach, endogeneity is an important issue, too. Therefore, we follow the same procedure described above. Though we have an exact measure of trade openness, we do not have data on the number of firms. Further, bilateral freeness of trade depends on imports and exports, and thus the endogeneity issues are similar to those in the case of the trade-to-GDP ratio. For this reason we instrument Φ_i by a proxy calculated from the estimates of bilateral trade freeness in a gravity approach. Because the GDP is the weight in the measure and because this might be endogenous, we use population shares to approximate the weights in the constructed freeness index.

We calculate bilateral trade openness, following Head and Ries (2001), according to

$$\phi_{ij} = \sqrt{\frac{I_{ji}}{Y_i - E_{ij}} \frac{I_{ij}}{Y_j - E_{ji}}}. \quad (17)$$

The correlation between the two measures is 0.3679. **Figure 1** provides a scatterplot of trade-

Table 1: Population-weighted coefficient of variation of regional GDP p.c. based on Gennaioli et al. (2013)

High-Income Countries		Upper-Middle-Income Countries		Lower-Middle-Income Countries		Low Income Countries	
(GNI per capita \$12,476 or more)	<i>I</i>	(GNI per capita \$4,036 to \$12,475)	Country	(GNI per capita \$1,026 to \$4,035)	Country	(GNI per capita \$1,025 or less)	Country
Argentina	0.73	Azerbaijan	0.12	Albania	0.50	Benin	0.50
Australia	0.09	Brazil	0.48	Armenia	0.19	Burkina Faso	0.13
Austria	0.22	Chile	0.36	Belize	0.26	Cambodia	0.38
Belgium	0.36	China	0.54	Bolivia	0.36	Cameroon	0.30
Bosnia and Herz.	0.09	Colombia	0.38	Egypt	0.33	Congo, Dem. Rep.	0.85
Bulgaria	0.24	Dominican Rep.	0.32	El Salvador	0.28	Ghana	0.38
Canada	0.21	Ecuador	0.29	Georgia	0.19	Kenya	0.91
Croatia	0.34	Gabon	0.24	Guatemala	0.47	Kyrgyz Republic	0.49
Czech Republic	0.40	Iran	0.63	Honduras	0.44	Madagascar	0.32
Denmark	0.24	Jordan	0.20	India	0.44	Malawi	0.12
Estonia	0.42	Kazakhstan	0.30	Indonesia	0.87	Mozambique	1.01
Finland	0.16	Latvia	0.59	Lao PDR	0.32	Nepal	0.17
France	0.26	Lebanon	0.24	Lesotho	0.35	Niger	0.42
Germany	0.20	Lithuania	0.31	Moldova	0.29	Tanzania	0.35
Greece	0.29	Macedonia, FYR	0.45	Mongolia	0.57	Uganda	0.30
Hungary	0.41	Malaysia	0.38	Morocco	0.23	Zimbabwe	0.32
Ireland	0.17	Mexico	0.58	Nicaragua	0.36		
Israel	0.12	Namibia	0.72	Nigeria	0.43		
Italy	0.26	Panama	0.48	Pakistan	0.08		
Japan	0.30	Peru	0.47	Paraguay	0.29		
Korea, Rep.	0.22	Romania	0.40	Philippines	0.25		
Netherlands	0.14	Russia	0.93	Senegal	0.16		
New Zealand	0.19	Serbia	0.39	Sri Lanka	0.50		
Norway	0.35	South Africa	0.44	Swaziland	0.21		
Poland	0.26	Thailand	0.87	Syrian Arab Rep.	0.11		
Portugal	0.26	Turkey	0.34	Ukraine	0.57		
Slovak Republic	0.53	Uruguay	0.37	Uzbekistan	0.43		
Slovenia	0.26	Venezuela	0.29	Vietnam	0.33		
Spain	0.20			Zambia	0.34		
Sweden	0.21						
Switzerland	0.27						
United Arab Em.	0.44						
United Kingdom	0.22						
United States	0.17						
Mean (unweighted)	0.27		0.43		0.35		0.43
Standard deviation	0.13		0.19		0.16		0.27

Figure 1: Trade-to-GDP ratio versus trade freeness (phiness)

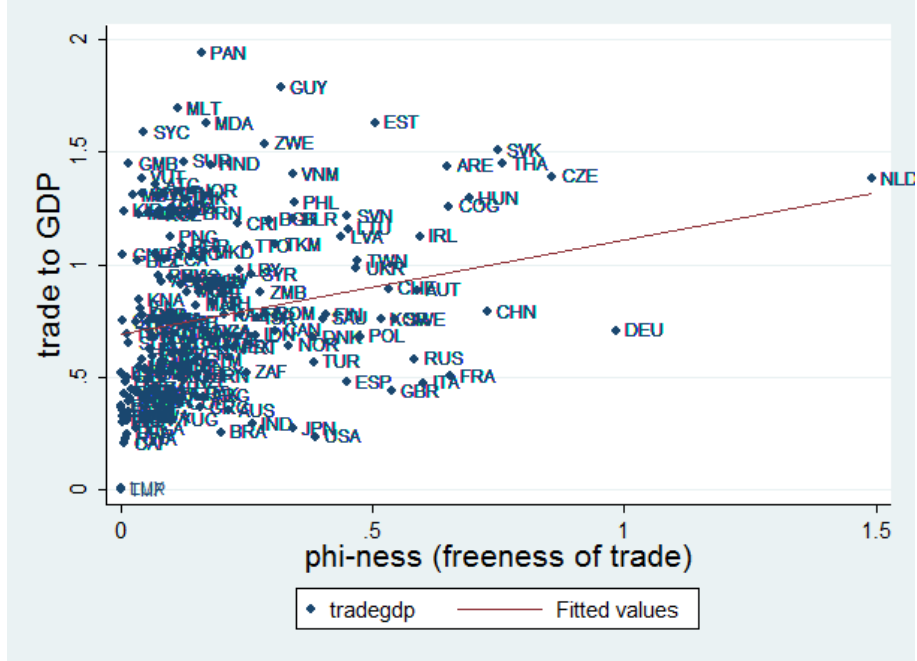


Table 2: Correlation of openness measures

	T_i	Φ_i	DO_i
T_i	1.0000		
Φ_i	0.3024	1.0000	
DO_i	0.4860	0.3309	1.0000

DO_i : degree of openness á la Arrabias et al. (2009).

Observations: 489589.

to-GDP ratio and freeness of trade, showing that larger countries (e.g., the USA, China, Germany) tend to have a higher freeness measure in comparison with their trade-to-GDP ratio.

Third, we use the openness measure provided by Arrabias et al. (2009), based on the ratio of exports to standardized GDP, as a third measure of openness, to control for robustness. Due to endogeneity, we use this measure only in the system GMM estimates. The correlation among these measures is displayed in **Table 2**.

3.1.3 Weights and interregional accessibility

As the theory shows (see (9) and (13)), interregional allocation matters, too. It determines the weights in the trade-to-GDP ratio as well as the nonlinearity in the freeness of trade, shown in the price indices that are a component of local and interregional market access. We approximate

this through interaction terms in the econometric specification – see (10) and (15) – and also by applying semi-parametric estimates.

Due to the lack of data on interregional accessibility, the number of firms, interregional trade, and trade between regions and other countries, we have to choose proxies for the internal structure of a country. We use the degree of urbanization as proxy for the distribution of agglomeration effects, firms, and population, and the share of agriculture as proxy for the importance of industry (which determines the agglomeration–periphery pattern). As a third alternative measure, we use the road density – measured by the ratio of road kilometers to country size – which refers to infrastructure and, indirectly, to accessibility within the country¹¹. These controls are also used as proxies for interregional and local market access if required.

In the cross-section estimations, we control for the number of sub-national units within countries (in logs), country size (ln of area in square kilometers), and the ratio of these two variables to control for average unit size. These variables are used as proxies for the internal structure of a country (see, e.g., Lessmann, 2014, for details).

3.2 Other controls

To allow for Kuznets-curve effects (see above), we control for income by the p.c. GDP and its square.

Following Rodríguez-Pose (2012) and Lessmann (2014), we further control for policy parameters: government size (measured by the expenditure-to-GDP ratio), a federal dummy, and EU membership. There is a quite large literature on the effect of decentralization on regional inequality, which shows that federal countries have lower regional inequalities, particularly if high-income countries are considered (e.g., Rodríguez-Pose and Ezcurra, 2010, and Lessmann, 2012). The EU membership dummy and the government size measure should proxy for redistributive politics within countries, performed by international donors or the national government. However, the EU dummy also controls for a high level of internal integration specific to a selection of countries. Table A.2 in the appendix provides a detailed description and source information for all variables considered in the analysis. **Table 3** shows descriptive statistics.

¹¹In this, we deviate from Rodríguez-Pose (2012), who uses a joint index of rail and road density. Since data on railways are lacking for many countries, we decided to focus on roads solely.

Table 3: Descriptive Statistics (Gennaioli et al., 2013)

Variable	Obs	Mean	Std.Dev.	Min	Max
covwshleifer	105	0.3606	0.1910	0.0839	1.0123
trade _{gd} p	105	0.8077	0.3740	0.2190	2.0520
trade _{gd} p	105	0.6259	0.3431	0.1534	1.6930
Φ	105	0.2543	0.2931	0.0016	2.1583
$\hat{\Phi}$	105	0.1987	0.2302	0.0013	1.3148
logunits	105	2.4291	0.6395	0.6931	4.3820
logarea	105	12.4143	1.5688	9.2301	16.6128
log(area/units)	105	5.4930	1.8261	3.1572	16.0722
urban(ization)	105	56.2137	21.6385	12.2600	97.1800
roaddens	83	68.2550	84.9324	1	490.3333
agri	103	14.3068	12.3885	0.8836	51.8000

3.3 Constructing the instrument for trade shares

The major problem for our econometric analysis is that trade shares might be endogenously determined by regional inequality. Consider for example a shock in regional inequality, caused by the discovery of natural resources or a boom in a single industry of one region of a country. It is obvious that the shock will also affect the trade statistics of the country. Therefore trade and regional inequality are simultaneously determined. Other sources of endogeneity are omitted variables and/or measurement errors in variables. For example, regional inequality might cause political instability (see Buhaug et al., 2012, Deiwiks et al., 2012, and Lessmann, 2013, for empirical evidence). This will discourage foreign and domestic investments, increase protection, and reduce income, thereby reducing international trade and freeness of trade. For this reason we construct an instrument for trade shares and use it in instrumental variable regressions.

We derive the gravity equation for the bilateral trade-to-GDP ratio from our model (see Appendix B):

$$\ln \frac{I_{ji} + E_{ij}}{Y_j} = \beta_0 + \beta_1 \ln Y_j + \beta_2 \ln Y_i + \beta_3 \ln \phi_{ij} + \beta_4 \ln \left(\frac{Q_j^{\sigma-1}}{\Pi_i} + \frac{P_i^{\sigma-1}}{\Pi_j} \right) + \varepsilon_{ij}. \quad (18)$$

Bilateral trade between country i and country j depends on both countries' incomes (Y_i , Y_j), on trade openness between the two countries, ϕ_{ij} , and on two terms describing the attractiveness of the destination country in comparison with the rest of the world and the relative market access to all countries. The latter is the multilateral resistance (ML) term Π of the country of origin. It

has to be considered when estimating bilateral trade (Anderson and van Wincoop, 2003)¹².

3.3.1 The bilateral trade equation

When estimating a gravity equation five problems arise: There might be (1) endogeneity of the GDP (Frankel and Romer, 1999) and (2) endogeneity of trade cost controls such as free trade agreements (e.g., Baier and Bergstrom, 2007). (3) One has to take account of multilateral resistance terms. (4) A huge number of zero trade flows are present in the data. (5) There will be unobserved heterogeneity across countries.

To handle endogeneity referring to the GDP we decide to follow Frankel and Romer (1999)¹³. However, we need a time-varying instrument of trade flows and hence extend their approach according to Head et al. (2010). For this, we use the population size of the countries and other time-invariant geographic variables (e.g., distance and access to the sea) instead of the endogenous GDP as determinants of bilateral trade.

To consider the multilateral resistance term (Anderson and van Wincoop, 2003) we follow Feenstra (2004) and include country-pair fixed effects. These dummy variables capture all time-invariant factors that are country-pair-specific (distance, common language, etc.)¹⁴. In doing so we also deal with endogeneity of trade-openness variables that affect trade, such as regional trade agreements (see Baier and Bergstrand, 2004). We further use distance and the interaction term in time and distance as additional proxies for trade openness.

Trade costs and a country's price index depend also on some intra-country variables such as the degree of agglomeration. However, because interregional inequality also depends on the degree of agglomeration and intra-country trade costs, controls that are used as proxy for those variables are not used in the gravity equation. Otherwise the instrument would become closely correlated with interregional inequality. Instead we use bilateral country-pair fixed effects to allow for a basic level of such effects.

Further, we control for exogenous changes in transport costs by an interaction term in distance and time, and for the importance of common institutional settings by considering an interaction term in the common legislative tradition and population size or, alternatively, an interaction term

¹²A similar expression has been provided by Helpman and Krugman (1985) and Helpman (1987).

¹³There is a discussion that the endogeneity of GDP might not be a huge problem because the GDP depends on net exports which are only a small contributor to GDP (e.g., Baier and Bergstrom, 2007).

¹⁴Cheng and Wall (1999) and Egger and Pfaffermayr (2003) have shown that neglecting these country-pair fixed effects causes an estimation bias.

in the dummy of common history and the population size. This refers to market access costs and to exogenous restrictions due to institutional settings¹⁵.

To summarize: we apply a regression with country-pair fixed effects and time fixed effects for constructing the cross-section proxy for the real trade-to-GDP ratio (see Baier and Bergstrom, 2007). As Cheng and Wall (1999) and Egger and Pfaffermayr (2003) have shown, neglecting these country-pair fixed effects would cause an estimation bias.

Our final bilateral gravity model has the following form:

$$\begin{aligned} \ln\left(\frac{\tau_{ijt}}{GDP_{it}}\right) = & \alpha_{ij} + \beta_1 \ln POP_{it} + \beta_2 \ln POP_{jt} \\ & + \beta_4 (Comlang_{ij} \times \ln POP_{jt}) + \beta_3 (Comlang_{ij} \times \ln POP_{it}) \\ & + \beta_5 (Time_t \times Distance) + \mu_t + \epsilon_{ijt}, \end{aligned} \quad (19)$$

where τ_{ijt} is the bilateral trade between countries i and j at time t measured as exports plus imports, GDP_{it} is the income of country i at time t , and $Comlang_{ij}$ represents a common language, which is used to proxy close historical and cultural links between countries. Thereby, we concentrate on bilateral trade-to-GDP ratios (trade shares) following the approach of Frankel and Romer (1999), not on exports as in that of Head et al. (2010).

Country size is considered by including the natural logarithm of the population size of countries i and j ($POP_{i,j}$). Finally, α_{ij} are country-pair fixed effects, μ_t are time fixed effects, and ϵ_{ijt} is the error term. Note that time-invariant trade determinants, such as the country size measured by surface area, are not included in the empirical model, since the country-pair fixed effects are perfectly correlated with these variables.

3.3.2 Results

We use the same data set as Head et al. (2010)¹⁶. The original source of the trade data is the International Monetary Fund's Direction of Trade Statistics (DOTS). The final data set covers the 1948–2006 period and contains trade data on 208 countries.

The main estimation results of the gravity model are reported in **Table 4**¹⁷. Column (1)

¹⁵One might consider extractive versus inclusive institutions as suggested by Acemoglu, Johnson, and Robinson (Acemoglu and Robinson, 2000, Acemoglu et al., 2002).

¹⁶The data set can be downloaded at <http://www.cepii.fr>

¹⁷We used many other specifications, including that of a common legal system or a shared history instead of a common language. We also applied a Poisson pseudo-maximum-likelihood estimation with bilateral fixed effects to

Table 4: The bilateral trade equation

Variable	(1) logflow_GDP	(2) Phiness
logpop_o	0.0732* (0.0390)	-0.161*** (0.0372)
logpop_d	0.0375 (0.0363)	-0.409 (0.0372)
Comlang \times logpop_o	0.0542 (0.0746)	
Comlang \times logpop_d	-0.191** (0.0753)	
tdummy_dist	-0.00581*** (0.000470)	-0.00563*** (0.000475)
Constant	-5.214*** (0.248)	-4.440*** (0.259)
Observations	714,343	489,724
R^2	0.013	0.031
Number of country pairs	28,621	24,287
F	48.39	77.24
Country-pair fixed effects and time effects		
The dependent variables are $\ln(\tau_{ijt}/GDP_{it})$ in (1) and Φ_i in (2). Robust standard errors in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.		

of **Table 4** displays the independent variable, column (2) the estimated coefficients, column (3) the robust standard errors, and column (4) the significance levels. The regression is based on 714,343 observations of 28,621 trade pairs (208 countries). The figures in column (1) provide evidence that trade openness is increasing in the population size of country i , while there is no significant unconditional effect of the size of country j . Concerning the interaction variables, only the interaction of the population in country j and the border dummy is significant, with a positive coefficient. Thus, the population size of trade partner countries j has an effect only in neighboring countries, not in general.

As displayed in column (2), we find evidence that trade freeness of a country declines with increasing population and distance, while the population size of the destination country is insignificant.

The estimated coefficients are then used to make linear predictions of the bilateral trade flows.

allow for the large number of zeros (e.g., Santos Silva and Tenreiro, 2011). To allow for dyadic FE in that approach we applied xtpqml, written by T. Simcoe.

Note that we use the full data set for this exercise – not only those countries which are included in the database used below. If we calculate the pairwise correlation between predicted bilateral trade and the dependent variable $\ln(\tau_{ijt}/GDP_{it})$, we get a high correlation of 0.9060. This is, however, a correlation between logarithms at the country-pair level, which is not the variable that we use in the main part of the paper. The correlation between bilateral trade freeness and the predicted trade freeness is 0.7643.

3.3.3 Aggregate trade and quality of the instrument

The predictions of the bilateral trade flows have to be aggregated for each country i . In this, we closely follow Frankel and Romer (1999). First, we rewrite equation (19) as

$$\ln(\tau_{ijt}/GDP_{it}) = \alpha_{ij} + \hat{\beta}\mathbf{X}_{i,j,t} + \mu_t, \quad (20)$$

where $\hat{\beta}$ is the vector of the estimated coefficients as reported in Table 4, and \mathbf{X}_{ijt} is the vector of trade determinants. Country i 's overall constructed trade share is then given by

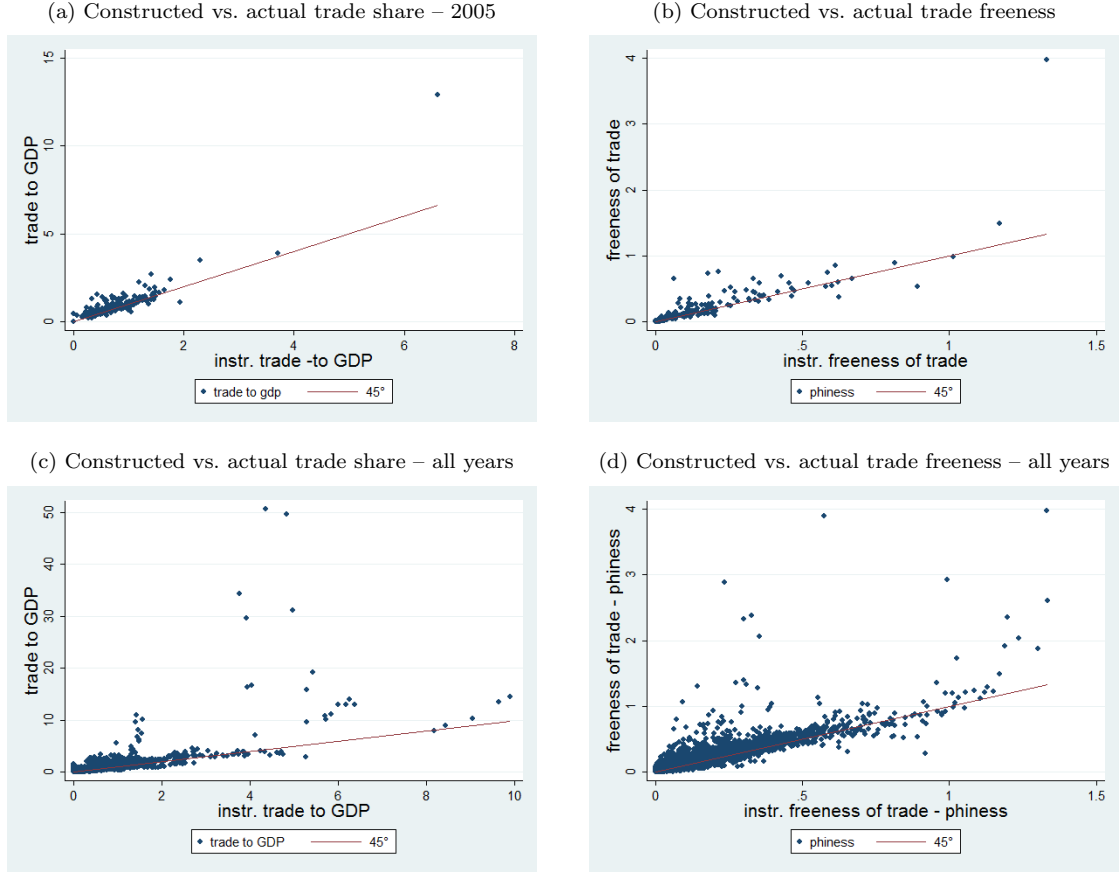
$$\hat{T}_{it} = \sum_{j \neq i} e^{\alpha_{ij} + \hat{\beta}\mathbf{X}_{i,j,t}}. \quad (21)$$

The constructed trade share of country i is given by the sum of the estimated bilateral trade with each country of the world. We are able to construct \hat{T}_{it} for 208 countries in the world. Note that the constructed trade flows are based only on the variables used in the gravity model; in particular, the trade shares are independent of country i 's income. We are also able to construct trade shares even for those countries where trade data is missing but trade determinants are available.

Figure 2 shows the relationship between the constructed openness measures, \hat{T}_{it} and $\hat{\Phi}$, and the actual openness measures: T_{it} as reported by the World Development Indicators Series, and Φ as calculated from the World Development Indicators Series. Panels (a) and (b) display it for the year 2005, and panels (c) and (d) for the whole time period considered.

The correlation between the constructed trade shares and actual trade shares is quite high, with a correlation coefficient of 0.7797. In the case of trade freeness the correlation is 0.8504. The figures also include the bisecting line, which indicates where actual and fitted openness measures are equal.

Figure 2: Constructed versus actual openness measures



3.4 IV regressions: cross-country results

In the following, we present the results of instrumental variable regressions using the cross-section data set. For this, the trade share is treated as endogenous, and the constructed trade shares from the gravity model are used as instruments. In order to reduce the effect of single outliers in the trade data, we build a 5-year period average of all variables (2000–2004). In the cross-section analysis, we refer to the data set based on the regional incomes used by Gennaioli et al. (2013), who compile the data for the year 2005. The final sample consists of 105 countries. **Table 5** presents the results of the trade-to-GDP regressions, and **Table 6** the second-stage results of the freeness-of-trade estimates¹⁸.

¹⁸Because OLS is very sensitive to outliers, we apply MM estimators as suggested by Yohai (1987) in the non-IV estimates. We applied the mmregress of Verardi and Croux (2009) to identify outliers with vertical and bad leverage and a robust estimator (robreg, see Jann 2010). As a robustness check we carried out IV estimates without outliers, but this did not change the findings.

Table 5: Cross section with trade-to-GDP ratio, Gennaioli's Data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Variables	IV0	IV1	IV2	IV3	IV4	IV5	IV9	IV10	RobR2
	covw	covw	covw	covw	covw	covw	covw	covw	covw
tradegdp	−0.0192 (0.0557)	0.209*** (0.0798)	0.229*** (0.0843)	0.199** (0.0862)	0.226*** (0.0856)	0.242*** (0.0904)	0.204** (0.0879)	0.211** (0.0941)	0.0899** (0.0422)
loggdppc	0.131 (0.162)	0.0167 (0.131)	0.0595 (0.156)	0.198* (0.113)	0.0573 (0.156)	0.142 (0.202)	0.120 (0.190)	0.0682 (0.189)	0.198* (0.113)
loggdppc2	−0.0104 (0.00984)	−0.0034 (0.00800)	−0.0052 (0.00907)	−0.0129* (0.00684)	−0.005 (0.00910)	−0.0094 (0.0111)	−0.009 (0.0110)	−0.0046 (0.0113)	−0.0136** (0.00663)
urban			−0.0014 (0.0015)	−0.0019 (0.0014)	−0.0015 (0.0015)	−0.0015 (0.0015)	−0.0017 (0.0016)	−0.0019 (0.0016)	−0.00053 (0.0012)
roaddens				−5.55e−05 (0.0002)					
popdens					−2.09e−05 (5.35e−05)				
agri						0.0024 (0.0038)			
govsize							−0.00087 (0.0022)	−0.00035 (0.0022)	
oecd								−0.0704 (0.0643)	
Constant, logunits, logarea, logarea/logunits									
Observations	105	105	105	83	105	103	83	83	105
R ²	0.088	0.304	0.304	0.329	0.306	0.293	0.387	0.390	
F	6.097	10.89	9.098	7.493	7.959	7.447	9.129	7.964	.
rss	3.460	2.640	2.641	1.589	2.634	2.636	1.883	1.874	.
idstat	20.36	10.61	11.32	9.239	11.06	12.11	9.131	8.408	.
idp	6.41e−06	0.0011	0.0008	0.0024	0.00089	0.0005	0.0025	0.0037	.
widstat	133.5	28.37	41.07	57.74	39.41	35.90	36.17	33.65	.

Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 6: Cross section with trade freeness, Gennaioli's data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	IVS0	IVS1	IV2	IV3	IV4	IV5	IV9	IV10	RobR2
	covw	covw	covw	covw	covw	covw	covw	covw	covw
phiness	0.0975 (0.0717)	0.0969* (0.0557)	0.103* (0.0583)	0.126 (0.108)	0.150*** (0.0583)	0.0984 (0.0617)	0.0971 (0.0599)	0.106* (0.0612)	0.0950*** (0.0271)
loggdppc	0.138 (0.160)	0.109 (0.137)	0.169 (0.171)	0.315*** (0.106)	0.161 (0.167)	0.144 (0.219)	0.210 (0.207)	0.165 (0.206)	0.262*** (0.101)
loggdppc2	−0.0115 (0.0099)	−0.00967 (0.0085)	−0.0124 (0.0101)	−0.0202*** (0.0063)	−0.0120 (0.0098)	−0.0112 (0.0124)	−0.0151 (0.0121)	−0.0112 (0.0123)	−0.0181*** (0.00600)
urban			−0.0017 (0.0014)	−0.0022 (0.0015)	−0.0018 (0.0014)	−0.0017 (0.0015)	−0.002 (0.0016)	−0.0023 (0.0016)	−0.0009 (0.00094)
roaddens				−0.00030 (0.0004)					
popdens					−0.0001** (5.29e−05)				
agri						−0.0010 (0.0032)			
govsize							−0.0002 (0.0019)	0.0003 (0.0018)	
oecd								−0.0666 (0.0669)	
Constant, logunits, logarea, logarea/logunits									
Observations	105	105	105	83	105	103	83	83	105
R ²	0.115	0.296	0.308	0.345	0.331	0.302	0.399	0.406	
F	5.909	8.853	7.229	6.796	7.215	5.918	8.452	7.433	.
rss	3.357	2.673	2.625	1.552	2.540	2.603	1.847	1.826	.
idstat	5.712	6.227	6.525	7.881	6.685	6.663	6.641	6.099	.
idp	0.0168	0.0126	0.0106	0.0050	0.0097	0.0098	0.010	0.0135	.
widstat	35.43	37.58	38.00	52.21	34.61	36.61	34.63	30.79	.

Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

In all IV estimates the first-round results provide a Kleibergen–Paap Wald statistic (widstat) that is always above 16, indicating that the instruments are not weak (see Stock and Yogo, 2005, for critical values). Further tests on endogeneity and underidentification (idstat) have the right signs, too.

Concerning the estimates of the trade-to-GDP ratio (**Table 5**), the point estimates imply that an increase in the ratio by 10 percentage points is associated with an increase in regional inequalities by approximately 2%. The regressions also imply that country size plays a role. In line with Lessmann (2014), we find evidence of an inverted-U relationship between development and regional inequalities. The effects of urbanization and infrastructure are not significant, but it is quite suggestive that the regression coefficients are negative. The share of agriculture is also not significantly related to regional inequality; nor are the political variables (government size, federal dummy, and EU dummy).

Table 6 presents the cross-section results concerning the trade freeness measure with the Gennaioli et al. (2013) data. The coefficient of the freeness measure is always positive but changes between being significant and insignificant. The finding is much less robust. This has clear implications on policy, as discussed below.

Next, we present similar estimations using a cross section of the regional data provided by Lessmann (2014). Here, our sample is significantly smaller, involving at most 54 countries. **Table 7** presents the results of the second-stage regressions.

Again, we find a positive and significant effect of trade on regional inequality. The effect is slightly larger, with regression coefficients between 0.22 and 0.32. Note that the country sample is biased towards high-income countries. As before, the coefficient of freeness of trade is positive, but its significance is not robust. We take this coincidence with the results of the Gennaioli et al. data as an indicator that the Lessmann data do not suffer from a strong selection bias and its panel results can be used to derive inference for the larger database.

3.5 IV regressions with panel data

A second source of a potential endogeneity bias in our regression analysis is unobserved heterogeneity between countries. There may be a huge number of unobserved variables, which could render our relationship between trade and inequality endogenous. These omitted variables can be

Table 7: Regional inequality and trade (cross-section, Lessmann (2014) data)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variable	IV1	IV2	IV3	IV4	IV5	IV9	IV10	RobR2
	covw	covw	covw	covw	covw	covw	covw	covw
tradegdp	0.316*** (0.0774)	0.341*** (0.0778)	0.331*** (0.0889)	0.338*** (0.0762)	0.301*** (0.0765)	0.356*** (0.0817)	0.353*** (0.0788)	0.235*** (0.0641)
loggdppc	0.253* (0.134)	0.412** (0.164)	0.249 (0.154)	0.375** (0.174)	0.0146 (0.274)	0.406* (0.210)	0.419** (0.204)	0.348* (0.210)
loggdppc2	-0.0183** (0.00823)	-0.0259*** (0.00963)	-0.0170* (0.00917)	-0.0238** (0.0101)	-0.00548 (0.0148)	-0.0253** (0.0123)	-0.0239** (0.0118)	-0.0232* (0.0125)
urban		-0.0032** (0.0015)	-0.0012 (0.0013)	-0.0033** (0.0016)	-0.0036** (0.0018)	-0.0037** (0.0015)	-0.0042*** (0.0016)	-0.0012 (0.0011)
roaddens			-0.0002 (0.0002)					
popdensity				-7.45e-05 (5.42e-05)				
agri					-0.0100** (0.0047)			
govsize						-0.0028 (0.0022)	-0.0019 (0.0021)	
oecd							-0.106* (0.0564)	
Constant, logunits, logarea, logarea/logunits								
Observations	53	53	46	53	52	49	49	54
R ²	0.636	0.672	0.682	0.681	0.684	0.679	0.697	
F	14.57	10.76	9.289	9.723	9.853	8.661	7.681	.
rss	0.739	0.666	0.530	0.647	0.617	0.628	0.594	.
idstat	15.43	16.75	14.96	16.80	15.29	15.81	15.90	.
idp	8.57e-05	4.27e-05	0.0001	4.16e-05	9.22e-05	6.99e-05	6.67e-05	.
widstat	123.7	109.2	93.37	106.1	93.04	99.64	92.01	.

Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 8: Regional inequality and trade freeness (cross section, Lessmann (2014) data)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variable	IVLEp0	IVLEp1	IVLEp2	IVLEp4	IVLEp5	IVLEp9	IVLEp10	RobRLEp2
	covw	covw	covw	covw	covw	covw	covw	covw
phiness	0.131** (0.0579)	0.0905 (0.0582)	0.0987* (0.0593)	0.114** (0.0556)	0.0364 (0.0784)	0.112* (0.0653)	0.119* (0.0687)	0.0885** (0.0369)
loggdppc	0.139 (0.203)	0.267 (0.185)	0.379 (0.237)	0.335 (0.249)	-0.302 (0.358)	0.263 (0.308)	0.276 (0.312)	0.298 (0.273)
loggdppc2	-0.0145 (0.0118)	-0.0206* (0.0109)	-0.0262* (0.0135)	-0.0238* (0.0141)	0.0095 (0.0195)	-0.0191 (0.0175)	-0.0182 (0.0176)	-0.0215 (0.0151)
urban			-0.0025 (0.002)	-0.0022 (0.002)	-0.0027 (0.0024)	-0.003 (0.0019)	-0.0034* (0.002)	-0.00018 (0.0014)
popdens				-9.51e-05 (6.66e-05)				
agri					-0.0172*** (0.0056)			
govsize						-0.0028 (0.00322)	-0.0022 (0.00320)	
oecd							-0.0786 (0.0813)	
Constant, logunits, logarea, logarea/logunits								
Observations	53	53	53	53	52	49	49	53
R ²	0.453	0.543	0.565	0.581	0.585	0.579	0.589	
F	12.68	11.90	12.82	10.74	11.68	13.53	12.83	.
rss	1.111	0.928	0.883	0.850	0.812	0.825	0.806	.
idstat	4.492	5.283	5.656	5.155	6.138	5.700	5.741	.
idp	0.0341	0.0215	0.0174	0.0232	0.0132	0.0170	0.0166	.
widstat	24.49	27.08	26.39	23.16	21.73	25.30	24.56	.

Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

a serious source of bias in our estimated parameters. In our empirical investigation it is very likely that an omitted-variable bias occurs, since regional inequalities are affected by several geographic and historical patterns we cannot measure satisfactorily. Using panel data, we can control for unobserved factors by including country fixed effects and time fixed effects.

We estimate four different specifications. The first model is a pooled robust regression with time fixed effects, the second is a two-way fixed-effect approach, the third is a panel IV approach with the constructed openness measures as instrument, and the last is a system GMM model (Arellano and Bover, 1995, and Blundell and Bond, 1998). The system GMM estimator was designed for small- T large- N panels, which is the case for our data set. Note that we only consider the time-varying control variables in the panel regressions. The measure of regional inequality comes from Lessmann (2014) and covers the period 1980–2009. We use 5-year-averaged data to take care of the gaps in the data.¹⁹ **Table 9** reports the results considering the trade-to-GDP ratio, and **Table 10** those for the freeness-of-trade measure.

In these estimates we find a significantly positive coefficient of trade. The standard tests suggest in all IV specifications that the trade-to-GDP ratio is endogenous and that our instruments are not weak. The findings are confirmed for other controls, too (see Table A.8 in the appendix). To sum up, the panel regressions support the findings from the cross-section analysis: trade increases interregional inequality within countries.

However, when we control for EU membership the coefficient becomes insignificant. This indicates that integration might offset the negative effect of trade on interregional inequality. Therefore we use freeness of trade, which is our measure of market access, to examine the effect of integration. The estimates do not provide significant coefficients of freeness of trade except for the pooled regression. This is evidence that an increase in freeness of trade, i.e., an increase in integration in the world markets, tends to be neutral with respect to interregional inequality²⁰.

3.6 Robustness tests

We performed many different robustness tests; some of them are presented below. We first test for nonlinearities, present results for different inequality measures, present findings for interaction

¹⁹Period 1: 1980–1984; period 2: 1985–1989; ...; period 6: 2005–2009.

²⁰Our findings in the pooled regression are in line with those of Ezcurra and Rodríguez-Pose (2013), who use the KOR index of globalization as the indicator for integration. However, when considering endogeneity, the significance vanishes in our study, whereas they do not take account of endogeneity issues.

Table 9: Regional inequality and trade (panel, Lessmann (2014) data)

Variable	(1) Pool2 covw	(2) Pool3 covw	(3) FE2 covw	(4) FE3 covw	(5) PanIV2 covw	(6) PanIV3 covw	(7) SysGMMT2 covw	(8) SysGMMT3 covw	(9) SysGMMT4 covw
tradegdp	0.136*** (0.0325)	0.237*** (0.0532)	0.139** (0.0537)	0.117** (0.0516)	0.385*** (0.100)	0.381*** (0.102)	0.0345** (0.0147)	0.108*** (0.036)	0.0388** (0.0169)
loggdppc	0.187*** (0.0707)	0.309* (0.179)	0.209 (0.135)	0.265** (0.121)	0.241** (0.107)	0.264** (0.108)	0.0086 (0.0468)	-0.186 (0.141)	-0.0639 (0.0593)
loggdppc2	-0.0147*** (0.0042)	-0.0209* (0.0109)	-0.0116 (0.0079)	-0.0150* (0.008)	-0.0172** (0.0073)	-0.0186** (0.0073)	-0.0013 (0.003)	0.0094 (0.0076)	0.0022 (0.0032)
urban		-0.00062 (0.001)		-0.0049* (0.0026)		-0.0018 (0.0024)		0.0031*** (0.0013)	0.0013 (0.00084)
roaddens		-7.86e-05 (0.0002)						-0.000211** (9.01e-05)	
L.covw							0.941*** (0.0760)	1.001*** (0.0741)	0.922*** (0.0611)
popdensity									-2.91e-05 (1.96e-05)
Constant	-0.489* (0.284)	-1.544* (0.823)	-31.90 (25.93)	-23.90 (24.74)	-52.84* (29.63)	-50.27* (30.15)	0.0350 (0.159)	0.635 (0.611)	0.313 (0.243)
Constant, logunits, logarea, logarea/logunits, time dummies									
Observations	200	75	200	200	194	194	149	63	146
r2_w	0.823	0.779	0.468	0.497	0.271	0.282			
ar1p							0.0402	.	0.0697
ar2p							0.235	.	0.297

Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 10: Regional inequality and openness (phiness, panel, Lessmann (2014))

Variable	(1) Pool2 covw	(2) Pool3 covw	(3) FE2 covw	(4) FE3 covw	(5) PanIV2 covw	(6) PanIV3 covw	(7) SyGMM2 covw	(8) SyGMM3 covw	(9) SyGMM4 covw
phiness	0.0855*** (0.0289)	0.134*** (0.0365)	0.0594 (0.0602)	0.0103 (0.0918)	0.0624 (0.126)	0.0369 (0.234)	-0.0012 (0.0192)	-0.0014 (0.0220)	0.0120 (0.0226)
loggdppc	0.318*** (0.0694)	0.179 (0.359)	0.259* (0.133)	-0.0159 (0.308)	0.260** (0.102)	0.0457 (0.612)	0.0422 (0.0527)	0.0423 (0.0896)	0.0144 (0.0573)
loggdppc2	-0.0217*** (0.004)	-0.0149 (0.019)	-0.0143 (0.009)	0.0061 (0.0234)	-0.0144* (0.008)	0.0016 (0.0445)	-0.0036 (0.0032)	-0.0035 (0.0048)	-0.0021 (0.0033)
urban	-0.0009 (0.0006)	0.0008 (0.0017)	-0.0057* (0.0029)	-0.0034 (0.0036)	-0.0057*** (0.0018)	-0.0035 (0.005)	0.0004 (0.0007)	0.0012 (0.0014)	0.0007 (0.0007)
roaddens		-0.0002 (0.0003)		-0.00015 (0.0003)		-0.00016 (0.0007)		-4.51e-05 (4.33e-05)	
L.covw							0.903*** (0.0421)	0.975*** (0.0459)	0.902*** (0.0406)
popdens									-2.21e-05* (1.19e-05)
Constant, logunits, logarea, logarea/logunits, time dummies									
Observations	194	73	194	73	194	73	144	61	141
r2_w	0.771	0.759	0.472	0.356	0.472	0.350			
ar1p							0.0198	.	0.0267
ar1							-2.331	.	-2.216
ar2p							0.259	.	0.306
ar2							-1.129	.	-1.023

Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

terms, and discuss endogeneity of the GDP variable. Here, we carry out estimates with lagged GDP. Further estimates that we do not present are robust non-IV regressions, which are close to the IV estimates presented above²¹.

3.6.1 Nonlinearities

Our theoretical model shows that effects might be nonlinear because they depend on the level of trade. Therefore, we perform semiparametric estimates and test the null hypothesis that the linear and nonlinear approaches do not significantly differ²². The equation to be estimated has the following form (we omit subscripts for clarity):

$$I = \alpha + f(T) + \gamma X + \epsilon, \quad (22)$$

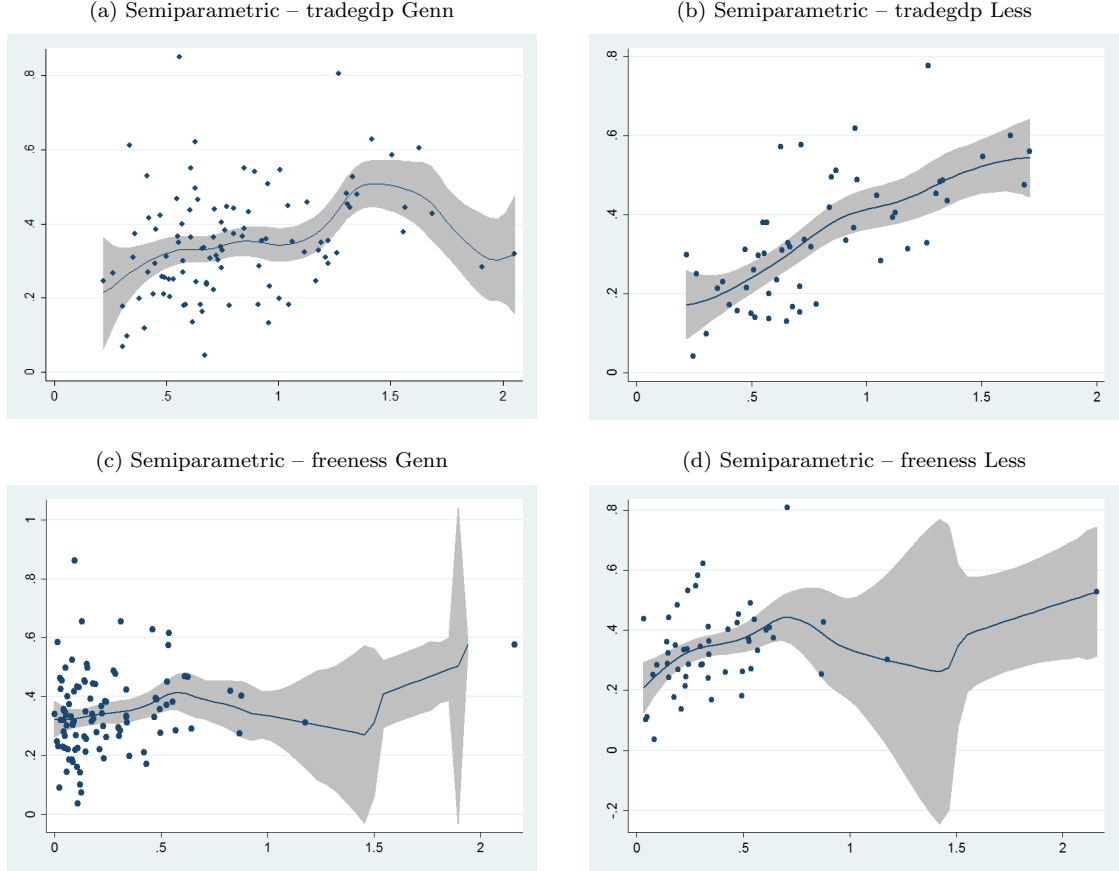
where X is a set of explanatory variables that are assumed to have a linear effect on the regional inequality I ; $f(\cdot)$ is an unknown smooth function of trade T , which might be nonlinear; and ϵ is a random error term. Thus, γX represents the parametric and $f(Y)$ the nonparametric part of the model. We refer to the estimator proposed by Robinson (1988). The intuition for this estimator is the following: in a first step, an estimate of $\hat{\gamma}$ is obtained using a procedure that is similar to the way in which variables can be partialled out of an OLS regression (but using nonparametric regressions); in a second step, a kernel regression of $I - \hat{\gamma}X$ on T is performed. In all stages, a Gaussian-kernel-weighted local polynomial fit is used for kernel regressions. This estimation procedure has been implemented in Stata by Verardi and Debarsy (2012). The results are presented in **Figure 3**.

Importantly, there is no sign of a systematic nonlinearity in the trade–inequality relationship. This is confirmed by the Hardle–Mammen (1993) statistic, which does not reject the null hypothesis that parametric linear and non-parametric fits are not different (p -values, from upper left to lower right: 0.15, 0.74; 0.52, 0.29).

²¹We even identified outliers with the `mmregress` procedure in Stata for each IV estimate, eliminated the outliers, and, carried out the same IV estimates as presented above. This did not change the results, so that we decided to present the results with the full sample. This implied nicer pictures for the semiparametric estimates presented below, which, however, we do not refer to. The results are not affected.

²²For instance, DiNardo and Tobias (2001) provide a discussion of semiparametric methods.

Figure 3: Semiparametric estimates. Plots of the estimated partial-regression functions. The inequality measure has been adjusted for the effects of the linear explanatory variables in the model [see equation (6)]. Shaded areas correspond to 90% confidence bands.



3.6.2 Alternative inequality measures

Next we consider alternative measures of regional inequality in the cross-section data set as well as in the panel. Different inequality measures usually do not provide an unambiguous country ranking. For example, the Gini coefficient is most sensitive to inequalities in the middle part of a distribution, while the coefficient of variation is strongly influenced by anomalously low or high income values. Moreover, the inequality measure used in the main part of the paper weights the income of sub-national regions by their population share. One can therefore interpret this measure as an intergroup inequality measure, where the groups of the population are defined by their location. If one is interested in purely geographic inequality, one should omit the population weights. We have therefore also considered unweighted inequality measures. For this purpose

Table 11: Alternative inequality measures - correlations

	Weighted coefficient of variation	Gini coefficient	Coefficient of variation
Weighted coefficient of variation	1.0000		
Gini coefficient	0.8997	1.0000	
Coefficient of variation	0.8848	0.9383	1.0000
Notes: Pairwise correlation between different inequality measures calculated based on Gennaioli et al. (2013).			

we calculate the Gini coefficient and the (unweighted) coefficient of variation of regional incomes. Table 11 shows pairwise correlation coefficients between the different inequality measures, based on the data provided by Gennaioli et al. (2013).

The correlation is around 0.9, suggesting that the particular measurement of regional inequality does not make much of a difference in the regression analysis.

Nevertheless, we check whether our findings are robust to the consideration of alternative inequality measures. For this purpose we run IV regressions in the large-cross-section and system GMM estimations in the panel. The results are reported in **Table 12**.

We report IV regressions in columns (1) and (2) and panel regressions in columns (3)–(5). Using the Gini coefficient (columns (3) and (5)) yields similar effects to the results reported in the previous sections (columns (1) and (3)). We find a positive and significant effect of trade on regional inequality. Using the unweighted coefficient of variation produces a positive – but insignificant – effect in the panel (see column (4)).

3.6.3 Lagged income

One may reasonably argue that income is also endogenous, since regional inequality might have a negative feedback effect on income. Therefore, we have repeated our cross-section estimates using historical income data instead of contemporaneous values. The period where we observe the per capita GDP refers to 1990–1994, while all other variables in the model refer to the period 2000–2004 (inequality refers to 2005). The results are reported in **Table A.13** in the appendix and confirm our finding of a positive effect of trade on regional inequality.

Table 12: Alternative inequality measures

Instrument: trade-to-GDP ratio – GINI COV					
Variable	(1) Rr1 covw	(2) Rr2 adgini	(3) Rr3 covw	(4) Rr4 cov	(5) Rr5 adgini
L.covw			0.923*** (0.0454)		
trade_gdp	0.329*** (0.0449)	0.212*** (0.0247)	0.0464** (0.0206)	0.0481 (0.0291)	0.0315*** (0.0103)
loggdppc	0.341*** (0.0824)	0.215*** (0.0443)	−0.0508 (0.0651)	−0.0721 (0.0639)	−0.0256 (0.0307)
loggdppc2	−0.0221*** (0.00484)	−0.0137*** (0.00258)	0.00149 (0.00360)	0.00281 (0.00354)	0.000878 (0.00171)
urban	−0.00268*** (0.000781)	−0.00147*** (0.000375)	0.00125 (0.000839)	0.000456 (0.000958)	0.000192 (0.000418)
logunits	0.139*** (0.0234)	0.0644*** (0.0123)			
logarea	0.0324*** (0.00772)	0.0261*** (0.00347)			
logarea/logunits	0.00226 (0.00503)	0.00616** (0.00264)			
L.cov				0.905*** (0.0535)	
L.adgini					0.879*** (0.0414)
Constant	−1.781*** (0.362)	−1.218*** (0.198)	0.248 (0.233)	0.395* (0.226)	0.152 (0.110)
Time dummies					
Observations	194	198	149	148	151
R^2	0.665	0.610			
F	37.01	37.19	151.7	170.4	194.6
ar1p			0.0328	0.107	0.0884
ar2p			0.223	0.971	0.273

Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

3.6.4 Interactions

A final set of robustness tests is related to interaction terms that are important for the trade-to-GDP ratio as shown in the theory. Because interaction terms with trade are endogenous, we build instruments with the exogenous variables (e.g., the degree of urbanization) and our proxies and perform IV estimates.

Table 13 presents the results for the IV regressions with interaction terms. The tests indicate that endogeneity cannot be rejected and that the instruments might not be weak. Despite that, the results are hard to interpret. In general, problems of multicollinearity are present. The variance inflation factor is high for all variables. This implies high standard errors and thus might explain the insignificant coefficients of the trade-to-GDP ratio. In different cases coefficients of the interaction term and of the variable used in the interaction term are of the same size and have opposite signs, though they are highly correlated. This is in particular true for *oecd*, *eu*, and *govsize*. Further, the degree of urbanization might also be endogenous, implying problems with the instruments. Nonetheless, the results indicate that interaction terms should be taken into account when one wants to examine the mechanism of the effects of openness on interregional inequality. The same is true with respect to *phi*-ness of trade (see **Table 14**). This is left to future research.

4 Summary and Conclusions

This paper studies the effects of international trade and openness on inequalities in interregional GDP per capita within countries. We derive two measures for international openness from a simple New Economic Geography model. One is the usually used trade-to-GDP ratio. The other is an aggregate freeness-of-trade measure that we propose and derive from aggregating bilateral freeness of trade (*phiness*). Further, we use two recent data sets on sub-national regions of countries, which allow us to calculate (time-varying) measures of regional inequality for a wide range of countries, including many developing countries. Up to 105 countries are considered in our cross-section data set; our panel data set consists of 54 countries for the period 1980–2009.

We apply gravity regressions to construct the proxies of the trade measures as instruments in the cross-section regressions. The IV regressions suggest a positive and significant effect of trade as well as of freeness of trade on regional inequality. The point estimate implies that an increase in the trade-to-GDP ratio by 10 percentage points is associated with an increase in regional inequalities

Table 13: Interaction effects (cross section; Gennaioli et al.'s data)

Variable	(1) IV3 covw	(2) IV4 covw	(3) IV5 covw	(4) IV6 covw	(5) IV7 covw	(6) IV8 covw	(7) IV9 covw
tradegdp	0.400 (0.346)	0.195 (0.135)	0.179* (0.102)	0.231** (0.0942)	0.0521 (0.208)	0.162 (0.111)	0.147 (0.114)
loggdppc	0.00767 (0.191)	0.197* (0.114)	0.0800 (0.162)	0.138 (0.200)	0.150 (0.188)	0.0184 (0.151)	0.133 (0.186)
loggdppc2	-0.0021 (0.0111)	-0.0128* (0.0068)	-0.0063 (0.0094)	-0.0091 (0.0109)	-0.0105 (0.0109)	-0.0006 (0.0089)	-0.0095 (0.0107)
urban	0.00057 (0.0039)	-0.0019 (0.0014)	-0.0017 (0.0015)	-0.0015 (0.0016)	-0.0018 (0.0016)	-0.0021 (0.0014)	-0.0021 (0.0017)
roaddens		-6.67e-05 (0.0003)					
popdens			-0.00017 (0.00017)				
agri				0.0011 (0.0054)			
govsize					-0.0059 (0.0058)		-0.0004 (0.0022)
oecd						-0.200* (0.110)	
eu							-0.119 (0.112)
trade \times urban	-0.0023 (0.0041)						
trade \times roaddens		0.001 (0.0225)					
trade \times popdens			0.0002 (0.0002)				
trade \times agri				0.00208 (0.00646)			
trade \times govsize					0.00569 (0.0062)		
trade \times oecd						0.127 (0.103)	
trade \times eu							0.127 (0.116)
Constant, logunits, logarea, logarea/logunits, time dummies							
Observations	105	83	105	103	83	105	83
R^2	0.261	0.330	0.301	0.280	0.396	0.354	0.419
F	7.789	6.665	7.169	6.451	8.696	8.093	8.337
rss	2.806	1.589	2.652	2.685	1.856	2.452	1.786
idstat	4.071	8.788	11.40	10.02	10.27	5.526	3.670
idp	0.0436	0.003	0.0007	0.0015	0.0014	0.0187	0.0554
widstat	2.467	31.65	20.30	14.89	14.54	8.681	7.501

Robust standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 14: Interaction effects freeness of trade (cross-section [38] data)

Variable	(1) IV1 covw	(2) IV2 covw	(3) IV3 covw	(4) IV4 covw	(5) IV5 covw	(6) IV6 covw	(7) IV7 covw
phiness	0.371 (0.315)	0.335 (0.399)	0.340 (0.328)	0.236** (0.104)	0.0218 (0.0714)	−0.107 (0.169)	0.416** (0.175)
loggdppc	0.117 (0.182)	0.304*** (0.115)	0.284** (0.111)	0.142 (0.165)	0.0644 (0.219)	0.116 (0.198)	0.0666 (0.212)
loggdppc2	−0.0095 (0.0105)	−0.0199*** (0.0065)	−0.0189*** (0.0065)	−0.0111 (0.0096)	−0.0059 (0.0124)	−0.0076 (0.0120)	−0.0057 (0.0126)
urban	−0.000921 (0.00173)	−0.00193 (0.00162)	−0.00212 (0.00144)	−0.00163 (0.00145)	−0.00148 (0.00150)	−0.00164 (0.00177)	−0.00181 (0.00161)
roaddens		−2.89e−05 (0.0005)					
popdens			−0.00012* (7.08e−05)	−7.72e−05 (6.49e−05)			
agri					−0.0023 (0.0031)		
govsize						−0.0122 (0.0099)	0.0006 (0.0018)
oecd							0.0186 (0.0838)
$\Phi \times \text{urban}$	−0.00329 (0.00354)						
$\Phi \times \text{roaddens}$		−0.0448 (0.0777)	−0.0365 (0.0525)				
$\Phi \times \text{popdens}$				−0.00015 (0.00012)			
$\Phi \times \text{agri}$					0.0224* (0.0124)		
$\Phi \times \text{govsize}$						0.0127 (0.0105)	
$\Phi \times \text{oecd}$							−0.378** (0.189)
Constant, logunits, logarea, logarea/logunits							
Observations	105	83	83	105	103	83	83
R^2	0.323	0.344	0.362	0.342	0.299	0.292	0.416
F	6.504	6.783	6.761	8.640	5.056	6.697	7.352
rss	2.571	1.554	1.511	2.497	2.616	2.175	1.794
idstat	24.35	7.480	12.53	12.35	13.87	2.217	12.15
idp	8.03e−07	0.00624	0.000401	0.000440	0.000196	0.136	0.000492
widstat	46.07	18.90	30.61	28.21	64.29	0.901	34.11

Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

by approximately 2%. Our findings for the trade-to-GDP ratio are confirmed for the panel of countries. In contrast, the panel approaches provided only insignificant coefficients of the freeness-of-trade measure, indicating that integration neutralizes the negative effect of trade on inequality.

What do we learn from this study? If countries increase trade, interregional inequalities increase, too. We found that interregional inequalities increase on average if countries increase their trade-to-GDP ratio. This outcome is fairly robust and supports earlier findings. Consequently, we have a classical trade-off between efficiency and distribution: while incomes rise in response to increased trade, inequalities rise, too. However, when we turn to openness, the results are ambiguous. Because the freeness-of-trade measure is an indicator for integration in the world markets, we interpret this as evidence that strengthening integration does on average neutralize the negative effect of trade on interregional inequality. We expect that this might also hold for other issues, such as wage inequality within countries and productivity effects of trade. Therefore, future research on those issues should also consider applying an aggregate openness measure like the one we propose.

References

- [1] Acemoglu, D., Johnson, S., Robinson, J.A., 2000. The colonial origins of comparative development: an empirical investigation. NBER Working Paper w7771.
- [2] Acemoglu, D., Johnson, S., 2000. Reversal of fortune: geography and institutions in making of the modern world income distribution. Quarterly Journal of Economics 117, 1231–1294.
- [3] Alcalá, P., Ciccone, A., 2004. Trade and productivity. Quarterly Journal of Economics 119, 612–645.
- [4] Alesina, A., Perotti, R., 1996. Income distribution, political stability, and investment. European Economic Review 40, 1203–1228.
- [5] Alonso Villar, O., 1999. Spatial distribution of production and international trade: a note. Regional Science and Urban Economics 29, 371–380.
- [6] Anderson, E., 2005. Openness and inequality in development countries: a review of theory and recent evidence. World Development 33, 1045–1063.
- [7] Anderson, J.E., van Wincoop, E., 2003. Gravity with Gravitas: A Solution to the Border Puzzle. American Economic Review 93, 170–192.
- [8] Arellano, M., Bover, O., 1995. Another look at the instrumental variables estimation of error-components models. Journal of Econometrics 68, 29–51.
- [9] Arrabias, I., Pérez, F., Tortosa-Ausina, E., 2009. Measuring globalization of international trade: Theory and evidence. World Development 37, 127–145.
- [10] Baier, S.L., Bergstrand, J.H., 2007. Do free trade agreements actually increase members' international trade? Journal of International Economics 71, 72–95.
- [11] Baldwin, R., Forslid, R., Martin, P., Ottaviano, G., Robert-Nicoud, F., 2005. Economic Geography & Public Policy. Princeton University Press.

- [12] Barrios, S., Strobl, E., 2009. The dynamics of regional inequalities. Regional Science and Urban Economics 39, 575–591.
- [13] Behrens, C., Gaingné, C., Ottaviano, G.I.P., Thisse, J., 2007. Countries, regions and trade: On the welfare impacts of economic integration. European Economic Review 51, 1277–1301.
- [14] Blundell, R., Bond, S., 1998. Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics 87, 115–143.
- [15] Bosker, M., Garretsen, H., 2010. Trade costs in empirical New Economic Geography. Papers in Regional Science 89, 485–512.
- [16] Breau, S., Rigby, D.L., 2010. International trade and wage inequality in Canada. Journal of Economic Geography 10, 55–86.
- [17] Brühlhart, M., Carrere, C., Trionfetti, F., 2012. How wages and employment adjust to trade liberalisation: Quasi-experimental evidence from Austria. Journal of International Economics 86, 68–81.
- [18] Brühlhart, M., 2011. The spatial effects of trade openness: a survey. Review of World Economics 147, 59–83.
- [19] Brühlhart, M., Crozet, M., Koenig, P., 2004. Enlargement and the EU periphery: the impact of changing market potential. World Economy 27, 853–875.
- [20] Buhaug, H., Gleditsch, K., Holtermann, H., Østby, G., Tollefsen, A.F., 2012. It's the local economy, stupid! Geographic wealth dispersion and conflict outbreak location. Journal of Conflict Resolution 55, 814–840.
- [21] Cheng, I.-H., Wall, H.-J., 1999. Controlling for heterogeneity in gravity models of trade and integration. Working Paper 1999–010E, Federal Reserve Bank of St. Louis.
- [22] Chiquiar, D., 2005. Why Mexico's regional income convergence broke down. Journal of Development Economics 77, 257–275.
- [23] Chiquiar, D., 2008. Globalization, regional wage differentials and the Stolper–Samuelson theorem: Evidence from Mexico. Journal of International Economics 74, 70–93.
- [24] Combes, R.-P., Lafourcade, M., Mayer, T., 2005. The trade-creating effect of business and social networks: evidence from France. Journal of International Economics 66, 1–29.
- [25] Crozet, M., Koenig-Soubeyran, P., 2004. EU enlargement and the internal geography of countries. Journal of Comparative Economics 32, 265–278.
- [26] Dalton, H., 1920. The Measurement of the Inequality of Incomes. Economic Journal 30, 348–361.
- [27] Deiwiiks, C., Cederman, L., Gleditsch, K., 2012. Inequality and conflict in federations. Journal of Peace Research 49, 289–304.
- [28] DiNardo, J., Tobias, J. L., 2001. Nonparametric Density and Regression Estimation. Journal of Economic Perspectives 15, 11–28.
- [29] Egger, H., Kreickemeier, U., 2012. Fairness, trade, and inequality. Journal of International Economics 86, 184–196.
- [30] Egger, P., Huber, P., Pfaffermayr, M., 2005. A note on export openness and regional wage disparity in Central and Eastern Europe. Annals of Regional Science 39, 63–71.
- [31] Egger, P., Pfaffermayr, M., 2003. The proper panel econometric specification of the gravity equation: a three-way model with bilateral interaction effects. Empirical Economics 28, 571–580.
- [32] Elbers, C., Lanjouw, P., Mistiaen, J., Özler, B., Simler, I.R., 2005. Are Neighbours Equal? Estimating Local Inequality in Three Developing Countries. In R. Kanbur and A. Venables (eds.), *Spatial Inequality and Development*, New York: Oxford University Press, pp. 37–76.
- [33] Ezcurra, R., Rodríguez-Pose, A., 2013. Does economic globalization affect regional inequality? A cross-country analysis. World Development 52, 92–103.

- [34] Faber, B., 2007. Towards the spatial patterns of sectoral adjustments to trade liberalisation: The case of NAFTA in Mexico. Growth and Change 38, 567–594.
- [35] Feenstra, R. 2004. Advanced International Trade. Princeton University Press, Princeton.
- [36] Frankel, J.A., Romer, D., 1999. Does trade cause growth? American Economic Review 89, 379–399.
- [37] Frankel, J.A., Rose, A., 2002. An estimate of the effects of common currencies on trade and income. Quarterly Journal of Economics 117, 437–466.
- [38] Gennaioli, N., La Porta, R. L., de Silanes, F. L., Shleifer, A., 2013. Human Capital and Regional Development. Quarterly Journal of Economics, forthcoming.
- [39] Gonzalez Rivas, M., 2007. The effects of trade openness on regional inequality in Mexico. Annals of Regional Science 41, 545–561.
- [40] Haaparanta, P., 1998. Regional concentration, trade, and welfare. Regional Science and Urban Economics 28, 445–463.
- [41] Hardle W., Mammen, E., 1993. Comparing nonparametric versus parametric regression fits. Annals of Statistics 21, 1926–1947.
- [42] Hanson, G. H., 1997. Increasing returns, trade and the regional structure of wages. Economic Journal 107, 113–133.
- [43] Hanson, G. H., 1998. Regional adjustment to trade liberalisation. Regional Science and Urban Economics 28, 419–444.
- [44] Head, K., Mayer, T., 2006. Regional wage and employment responses to market potential in the EU. Regional Science and Urban Economics 36, 573–594.
- [45] Head, K., Mayer, T., Ries, J., 2010. The erosion of colonial trade linkages after independence. Journal of International Economics 81, 1–14.
- [46] Head, K., Ries, J., 2011. Increasing returns versus national product differentiation as an explanation for the pattern of U.S.-Canada trade. American Economic Review 91, 858–876.
- [47] Helpman, E., 1998. The size of regions, in: D. Pines, E. Sadka, I. Zilcha (eds.), Topics in Public Economics. Theoretical and Empirical Analysis. Cambridge: Cambridge University Press, 33–54.
- [48] Heston, A., Summers, R., Aten, B., 2012. Penn World Table Version 7.1. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.
- [49] Jann, B. (2010). robreg: Stata module providing robust regression estimators. Available from <http://ideas.repec.org/c/boc/bocode/s457114.html>.
- [50] Junius, K., 1999. Primacy and economic development: Bell shaped or parallel growth of cities? Journal of Economic Development 24, 1–22.
- [51] Kanbur, R., Venables, A., 2005. Spatial Inequality and Development. In: R. Kanbur and A. Venables (editors), Spatial Inequality and Development, 3–11. New York: Oxford University Press.
- [52] Kanbur, R., Zhang, X., 2005. Fifty years of regional inequality in China: A journey through central planning, reform and openness. Review of Development Economics 9, 87–106.
- [53] Kessler, A.S., Hansen, N.A., Lessmann, C., 2011. Interregional Redistribution and Mobility in Federations: A Positive Approach. Review of Economic Studies 78, 1345–1378.
- [54] Krugman, P., 1991. Increasing returns and economic geography. Journal of Political Economy 99, 483–499.
- [55] Krugman, P., Livas Elizondo, R., 1996. Trade policy and the third world metropolis. Journal of Development Economics 49, 137–150.
- [56] Kuznets, S., 1955. Economic Growth and Income Inequality. American Economic Review 45, 1–28.

- [57] Lessmann, C., 2014. Spatial Inequality and Development – Is there an Inverted-U Relationship? Journal of Development Economics 106, 35–51.
- [58] Lessmann, C., 2013. Regional inequality and internal conflict. CESifo Working Paper No. 4112, CESifo Group: Munich.
- [59] Lessmann C., 2012. Regional inequality and decentralization: an empirical analysis. Environment and Planning A 44, 1363–1388.
- [60] Milanovic B., 2005. Half a world: regional inequality in five great federations. Journal of the Asia Pacific Economy 10:408–445.
- [61] Monfort, P., Nicolini, R., 2000. Regional convergence and international integration. Journal of Urban Economics 48, 286–306.
- [62] Monfort, P., van Ypersele, T., 2003. Integration, regional agglomeration and international trade. (CEPR Discussion Paper #3752).
- [63] Østby, G., Nordås, R., Rød, J.K., 2009. Regional Inequalities and Civil Conflict in Sub-Saharan Africa. International Studies Quarterly 53, 301–324.
- [64] Paluzie, E., 2001. Trade policies and regional inequalities. Papers in Regional Science 80, 67–85.
- [65] Pavcnik, N., 2013. Globalization and within-country income inequality, in: M. Bacchetta and M. Jansen (eds.). Making Globalization Socially Sustainable, chap. 7, Geneva: ILO. (www.wto.org/english/res_e/booksp_e/glob_soc_sus_e_chap7_e.pdf)
- [66] Pernia, E. M., Quising, P. F., 2003. Trade openness and regional development in a developing country. Annals of Regional Science 37, 391–406.
- [67] Petrakos G., Rodríguez-Pose, A., Rovolis, A., 2005. Growth, integration, and regional disparities in the European Union. Environment and Planning A 37, 1837–1855.
- [68] Pflüger, M., 2004. A simple, analytically solvable, Chamberlinian agglomeration model. Regional Science and Urban Economics 34, 565–573.
- [69] Pflüger, M., Suedekum, J., 2008. Integration, agglomeration and welfare. Journal of Urban Economics 63, 544–566.
- [70] Pigou, A. C., 1912. Wealth and Welfare. London: Macmillan.
- [71] Redding, S., Sturm, D., 2008. The costs of remoteness: Evidence from German division and reunification. American Economic Review 98, 1766–1797.
- [72] Redding, S., Venables, A., 2004. Economic Geography and international inequality. International Economic Review 62, 53–82.
- [73] Rodríguez-Pose, A., 2012. Trade and Regional Inequality. Economic Geography 88, 109–136.
- [74] Rodríguez-Pose, A., N. Gil, 2006. How does trade affect regional disparities? World Development 34, 201–222.
- [75] Rodríguez-Pose, A., Ezcurra, R. 2010. Does Decentralization Matter for Regional Disparities? A Cross-Country Analysis. Journal of Economic Geography 10, 619–644.
- [76] Rodríguez-Pose, A., Sanchez-Reaza, J., 2005. Economic polarization through trade: Trade liberalization and regional growth in Mexico. In A. J. Venables and R. Kanbur (eds.), Spatial inequality and development. Oxford, UK: Oxford University Press.
- [77] Robinson, P. M., 1988. Root-N-Consistent Semiparametric Regression. Econometrica 56, 931–54.
- [78] Rodrik D., van Ypersele, T. 2001. Capital mobility, distributive conflict and international tax coordination. Journal of International Economics 54, 57–73.
- [79] Sachs, J. D. and A. Warner, 1995. Economic Reform and the Process of Global Integration. Brookings Papers on Economic Activity 1995, 1–118.
- [80] Sanchez-Reaza J., Rodríguez-Pose, A., 2002. The impact of trade liberalization on regional disparities in Mexico. Growth and Change 33, 72–90.

- [81] Santos Silva, J.M.C., Tenreyro, S., 2011. Further simulation evidence of the performance of the Poisson pseudo-maximum likelihood estimator. Economics Letters 112, 220–222.
- [82] Sen, A., 1973. On Economic Inequality. New York: Norton.
- [83] Stock, J.H., Yogo, M., 2005. Testing for Weak Instruments in Linear IV Regression. In D. W. K. Andrews and J. H. Stock (eds.), Identification and Inference for Econometric Models: A Festschrift in Honor of Thomas J. Rothenberg. Cambridge, UK: Cambridge University Press, 80–104.
- [84] Treisman, D. (2008). Decentralization dataset. <http://www.sscnet.ucla.edu/polisci/faculty/treisman/>.
- [85] Verardi, V., Courx, Ch., 2009. Robust regression in Stata. Stata Journal 9, 439–453.
- [86] Verardi, V., Debarsy, N., 2012. Robinson’s square root of N consistent semiparametric regression estimator in Stata. Stata Journal 12, 726–735.
- [87] Volpe Martincus, C., 2010. Spatial effects of trade policy: Evidence from Brazil. Journal of Regional Science 50, 541–569.
- [88] Williamson, J.G., 1965. Regional Inequality and the Process of National Development: A Description of Patterns. Economic Development and Cultural Change 13, 3–45.
- [89] Wood A., 2002. Globalization and wage inequalities. A synthesis of three theories. Weltwirtschaftliches Archiv 138, 54–82.
- [90] World Bank, 2011. World Development Indicators 2010. Washington, D.C.: World Bank.
- [91] Yemtsov, R., 2005. Quo Vadis? Inequality and Poverty Dynamics across Russian Regions. In R. Kanbur and A. Venables (eds.), Spatial Inequality and Development, New York: Oxford University Press, pp. 348–408.
- [92] Yohai, V.J. 1987. High breakdown-point and high efficiency robust estimates for regression. Annals of Statistics 15, 642–656.
- [93] Zhang, X., Zhang, K.H., 2003. How does globalization affect regional inequality within a developing country? Evidence from China. Journal of Development Studies 39, 47–67.

A Income inequality. Phiness of trade

We substitute the price indices (4) into the equation (6) for operating profits, and divide by P_r to get (11). Next, we calculate the difference

$$\frac{y_r - y_i}{y_r} = \left(\frac{e_{Er}}{p_r} - \frac{e_{Ar}}{p_r} \right) \frac{E_i}{Y_i}. \quad (23)$$

First, the exports of Home are

$$E_i = N^\sigma \mu_i \sum_j \left(\frac{P_j \alpha_j}{\Delta_j} \sum_s \lambda_s \phi_{sj} \right) = N^\sigma Y_i^d \sum_j \gamma_j \frac{\mu_i \phi_{ij}}{\mu_i \phi_{ij} + \mu_j} = N^\sigma Y_i^d \gamma_i \Phi_i,$$

where $\phi_{ij} = \sum_s \lambda_s \phi_{sj}$ and $\gamma_j = (P_j \alpha_j) / Y_i^d$ with Y_i^d the aggregate GDP of all destination countries of exports of Home, and Φ_i the openness of country i . Further, the export share of region r is

$$e_{Er} = \frac{N^\sigma Y_i^d \sum_j \frac{\gamma_j \mu_i \lambda_r \phi_{rj}}{\Delta_j}}{N^\sigma Y_i^d \Phi_i} = \frac{\sum_j \frac{\gamma_j \mu_i \lambda_r \phi_{rj}}{\Delta_j}}{\Phi_i} = \lambda_r \sum_j \frac{\phi_{rj}}{\phi_{ij}}. \quad (24)$$

Next we assume that intra-country transport costs are symmetric and identical across regions ($\psi = \psi_{rs}$) and that expenditures for manufacturing are also identical across regions. Accordingly, the absorption share of region r is

$$e_{Ar} = \frac{\frac{\lambda_r p_r}{\Delta_r} + \psi \frac{\lambda_s p_s}{\Delta_s}}{(1 + \psi) \sum_s \frac{\lambda_s p_s}{\Delta_s}}. \quad (25)$$

Putting (24) and (25) into (23) yields the relative income difference

$$\frac{y_r - y_i}{y_i} = \frac{N^\sigma}{p_r} \left(\lambda_r \sum_j \frac{\phi_{rj}}{\phi_{ij}} - \frac{\frac{\lambda_r p_r}{\Delta_r} + \psi \frac{\lambda_s p_s}{\Delta_s}}{(1 + \psi) \sum_s \frac{\lambda_s p_s}{\Delta_s}} \right) \Phi_i, \quad (26)$$

where γ_i is the GDP of the country of origin i relative to the aggregate GDPs of all destination countries. If we assume that $\phi_{rj} = \phi_{ij}$, we get (13).

B The theoretical gravity equation

We can derive a gravity equation for bilateral trade from our theoretical model. To simplify matters we derive the equation for manufacturing and assume identical firms in Home. The imports of country i from another country j are the sum of the nominal imports of the regions of country i , i.e.,

$$\begin{aligned}
 I_{ji} &= \sum_r n_j P_i T_{jr} q_j m_{jr} \\
 &= n_j q_j^{1-\sigma} (\alpha_r P_r T_{jr} Q_r^{\sigma-1} + a_s P_s T_{js} Q_s^{\sigma-1}) \\
 &= n_j q_j^{1-\sigma} Y_i (e_{Yr} T_{jr} Q_r^{\sigma-1} + e_{Ys} T_{js} Q_s^{\sigma-1}) \\
 &= n_j q_j^{1-\sigma} Y_i T_{ij} Q_i^{\sigma-1}.
 \end{aligned} \tag{27}$$

T_{ij} denotes bilateral trade costs, Q_i is the country's consolidated price index, and e_{Yr} is the income share of region r of country i . Next we can define a country's constraint for monopolistic goods. The income of a country j arising from monopolistic production depends on all sales of monopolistic goods to all other countries, including country j itself, i.e.,

$$Y_i = \sum_j A_{ji}. \tag{28}$$

Substituting I_{ji} (see (28)) for A_{ji} and rearranging yields

$$n_j p_j^{1-\sigma} = \frac{Y_j}{\Pi_j}, \tag{29}$$

with the income share s_Y and the multilateral resistance (ML) term

$$\Pi_j \equiv \sum_i Y_i \phi_{ji} Q_i^{\sigma-1}.$$

The ML term represents the relative accessibility of all regions for exports and thus the attractiveness of alternative export destinations (e.g., Anderson and van Wincoop, 2003). Exports of country j to another country i are defined analogously to (27):

$$E_{ji} = n_j p_j^{1-\sigma} Y_i \phi_{ij} Q_i^{\sigma-1}. \tag{30}$$

Putting in (30) yields

$$E_{ji} = \frac{Y_i}{\Pi_i} Y_j \phi_{ij} Q_j^{\sigma-1}. \quad (32)$$

Adding bilateral imports (27) and exports (27) yields, with symmetric bilateral trade freeness,

$$I_{ij} + E_{ji} = Y_j Y_i \phi_{ij} \left(\frac{Q_j^{\sigma-1}}{\Pi_i} + \frac{Q_i^{\sigma-1}}{\Pi_j} \right). \quad (33)$$

From this we get the econometric equation for the bilateral trade-to-GDP ratio

$$\ln \frac{I_{ji} + E_{ij}}{Y_j} = \beta_0 + \beta_1 \ln Y_j + \beta_2 \ln Y_i + \beta_3 \ln \phi_{ij} + \beta_4 \ln \left(\frac{Q_j^{\sigma-1}}{\Pi_i} + \frac{P_i^{\sigma-1}}{\Pi_j} \right) + \varepsilon. \quad (34)$$

Table A.3 reports the results of the first-stage regressions, **Table A.4** (lower part) the results of the second-stage regressions.

Table A.1: Summary statistics based on cross-section data

Variable	Obs.	Mean	Stand. Dev.	Min.	Max.
Regional inequality	106	0.359	0.191	0.084	1.012
Trade share	106	0.802	0.376	0.219	2.052
Constructed trade share	106	0.634	0.357	0.142	1.753
GDP p.c.	106	7.759	1.581	4.419	10.557
Spatial units	106	2.443	0.653	0.693	4.382
Area	106	12.448	1.600	9.230	16.613
Urbanization	106	56.436	21.657	12.260	97.180
Share agriculture	104	14.181	12.395	.884	51.800
Road density	84	3.527	1.274	0.000	6.195
Government size	84	25.371	9.853	8.335	46.100
Federal dummy	106	0.179	0.385	0.000	1.000
EU dummy	106	0.226	0.420	0.000	1.000

Table A.2: Data sources and definitions

Variable	Definition	Source
Regional inequality	Population-weighted coefficient of variation of regional GDP per capita.	Gennaioli et al. (2013) and Lessmann (2014)
GDP p.c.	GDP per capita in 2005 \$ prices.	World Bank (2011)
Spatial units	ln of the number of regions considered for the calculation of measures of regional inequality.	Gennaioli et al. (2013) and Lessmann (2014)
Area	ln of area in square kilometers.	World Bank (2011)
Trade share	Sum of imports and exports (total trade) as a share of the GDP.	World Bank (2011)
Constructed trade share	Trade share as estimated based on a gravity model of bilateral trade.	Head et al. (2010)
Urbanization	Share of urban-living population in total population.	World Bank (2011)
Share agriculture	Share of agricultural value added in total gross value added.	World Bank (2011)
Road density	Kilometers of paved roads divided by country size.	World Bank (2011)
Government size	Ratio of government expenditures to GDP.	World Bank (2011)
Federal dummy	Dummy for countries with a federal constitution.	Treisman (2008)

Table A.3: Regional inequality and trade (cross section, Gennaioli et al. (2013) data)

First-stage regression					
	(1)	(2)	(3)	(4)	(5)
Variable	IV1r	IV2r	IV3r	IV4r	IV5r
	tradeGDP	tradeGDP	tradeGDP	tradeGDP	tradeGDP
hatflowGDP	0.706*** (0.135)	0.746*** (0.117)	0.805*** (0.138)	0.698*** (0.134)	0.671*** (0.140)
loggdppcc	0.238 (0.181)	0.474** (0.220)	0.326 (0.212)	0.198 (0.150)	-0.163 (0.312)
loggdppcc2	-0.0166 (0.0109)	-0.0282** (0.0124)	-0.0216* (0.0129)	-0.0140 (0.0092)	0.00351 (0.0169)
urban		-0.0048* (0.0026)			
roaddens			-0.0001 (0.0003)		
popdenssh				-0.00014 (9.40e-05)	
agri					-0.0123 (0.0075)
Constant, logunits, logarea, logarea/logunits					
Observations	101	103	81	102	98
R^2	0.560	0.578	0.657	0.569	0.583
F	25.95	26.85	34.61	27.61	24.74
Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$					

Table A.4: Regional inequality and trade (cross section, Gennaioli et al. (2013) data)

	(1)	(2)	(3)	(4)	(5)	(6)
Variable	IV1	IV2	IVSh3	IV4	IV5	RobR2
	covw	covw	covw	covw	covw	covw
tradeGDP	0.163** (0.0686)	0.214** (0.0836)	0.129** (0.0628)	0.182** (0.0741)	0.171** (0.0728)	0.0899** (0.0422)
loggdppc	0.194** (0.0822)	0.289** (0.113)	0.184* (0.100)	0.0410 (0.134)	0.343*** (0.125)	0.198* (0.113)
loggdppc2	-0.0138*** (0.00514)	-0.0184*** (0.0067)	-0.0128** (0.0063)	-0.0045 (0.0083)	-0.0222*** (0.0072)	-0.0136** (0.0067)
urban		-0.0024* (0.0013)				-0.0005 (0.0012)
roaddens			-1.17e-05 (0.0002)			
popdens				-2.21e-05 (4.65e-05)		
agri					0.0026 (0.0027)	
Constant, logunits, logarea, logarea/logunits						
Observations	101	103	81	102	98	105
R^2	0.413	0.346	0.379	0.378	0.370	.
F	11.96	9.330	11.07	10.65	7.055	.
rss	1.480	2.042	1.080	1.729	1.655	.
Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$						

Table A.5: Pooled robust regression – trade-to-GDP ratio – Lessmann

Variable	(1) Pool2 covw	(2) Pool03 covw	(3) Pool4 covw	(4) Pool5 covw	(5) Pool6 covw	(6) Pool7 covw	(7) Pool8 covw
tradegdp	0.136*** (0.0325)	0.237*** (0.0532)	0.136*** (0.0407)	0.124*** (0.0374)	0.203** (0.0871)	0.456*** (0.0795)	0.219*** (0.0724)
loggdppc	0.187*** (0.0707)	0.309* (0.179)	0.0454 (0.334)	0.105 (0.0891)	0.333 (0.544)	0.673*** (0.174)	0.343 (0.301)
loggdppc2	−0.0147*** (0.0042)	−0.0209* (0.0109)	−0.0068 (0.0177)	−0.0108** (0.0049)	−0.0234 (0.0349)	−0.0360*** (0.0087)	−0.0231 (0.0171)
urban		−0.00062 (0.001)	−0.00078 (0.0006)	−0.0005 (0.0007)	−0.0012 (0.0009)	−0.0055*** (0.0017)	−0.0019 (0.0012)
roaddens		−7.86e−05 (0.0002)					
popdens			−0.00014 (0.0001)				
agri				−0.0041** (0.0017)			
govsize					0.002 (0.0096)	−0.0027 (0.0017)	0.0021 (0.0021)
oecd						−0.155** (0.0649)	
eu							−0.0466 (0.0502)
Constant, logunits, logarea, logarea/logunits, time dummies							
Observations	200	75	200	195	115	115	115
Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$							

Table A.6: FE regression – trade-to-GDP ratio – Lessmann

Variable	(1) FE2 covw	(2) FE3 covw	(3) FE4 covw	(4) FE5 covw	(5) FE6 covw	(6) FE7 covw	(7) FE8 covw
tradegdp	0.139** (0.0537)	0.117** (0.0516)	0.0447 (0.104)	0.110** (0.0503)	0.107* (0.0548)	0.0415 (0.0606)	0.0415 (0.0606)
loggdppc	0.209 (0.135)	0.265** (0.121)	−0.00441 (0.217)	0.384* (0.201)	0.229 (0.149)	0.450* (0.235)	0.450* (0.235)
loggdppc2	−0.0116 (0.0079)	−0.0150* (0.008)	0.0049 (0.0172)	−0.0229* (0.0132)	−0.0129 (0.009)	−0.0243 (0.0147)	−0.0243 (0.0147)
urban		−0.00487* (0.00260)	−0.00234 (0.00473)	−0.00468* (0.00260)	−0.00643* (0.00333)	−0.0103* (0.00544)	−0.0103* (0.00544)
roaddens			−0.000271 (0.000263)				
popdens				−0.000199 (0.000157)			
agri					−0.00179 (0.00383)		
govsize						−0.00332* (0.00174)	−0.00332* (0.00174)
Constant, logunits, logarea, logarea/logunits, FE, time dummies							
Observations	200	200	75	200	195	115	115
R^2	0.468	0.497	0.367	0.507	0.505	0.562	0.562
F	6.733	6.540	3430	8.867	6.566	5.171	5.171
Robust standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$							

Table A.7: Panel IV – trade-to-GDP ratio – Lessmann

Variable	(1) FEIV2 covw	(2) FEIV3 covw	(3) FEIV4 covw	(4) FEIV5 covw	(5) FEIV6 covw	(6) FEIV7 covw	(7) FEIV8 covw
tradegdp	0.385*** (0.100)	0.381*** (0.102)	−0.421 (1.266)	0.381*** (0.102)	0.219** (0.103)	0.596* (0.353)	0.596* (0.353)
loggdppc	0.241** (0.107)	0.264** (0.108)	−0.467 (1.369)	0.311** (0.145)	0.269** (0.118)	0.400 (0.267)	0.400 (0.267)
loggdppc2	−0.0172** (0.00733)	−0.0186** (0.00725)	0.0408 (0.104)	−0.0217** (0.00943)	−0.0167** (0.00749)	−0.0237 (0.0179)	−0.0237 (0.0179)
urban		−0.0018 (0.0024)	−0.0125 (0.0285)	−0.0017 (0.0024)	−0.0043* (0.0026)	−0.0006 (0.0093)	−0.0006 (0.0093)
roaddens			0.0008 (0.0031)				
popdens				−7.55e−05 (0.0002)			
agri					−0.00073 (0.0027)		
govsize						−0.0002 (0.0036)	−0.0002 (0.0036)
Constant, logunits, logarea, logarea/logunits, FE, time dummies							
Observations	194	194	73	194	193	113	113
rss	0.216	0.213	0.0227	0.213	0.170	0.135	0.135

Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.8: System GMM – trade-to-GDP ratio – constructed instrument not used

Variable	(1) SysGMM2 covw	(2) SysGMM3 covw	(3) SysGMM4 covw	(4) SysGMM6 covw	(5) SysGMM7 covw	(6) SysGMM8 covw	(7) SysGMM9 covw
L.covw	0.941*** (0.0760)	1.001*** (0.0741)	0.922*** (0.0611)	0.992*** (0.0642)	0.992*** (0.0642)	1.042*** (0.0804)	0.963*** (0.0746)
tradegdp	0.0345** (0.0147)	0.108*** (0.0360)	0.0388** (0.0169)	0.0295* (0.0175)	0.0295* (0.0175)	0.0358* (0.0212)	0.0190 (0.0158)
loggdppcc	0.0086 (0.0468)	−0.186 (0.141)	−0.0639 (0.0593)	0.0123 (0.0710)	0.0123 (0.0710)	0.0117 (0.0947)	−0.0498 (0.0818)
loggdppcc2	−0.0013 (0.003)	0.0094 (0.0076)	0.0022 (0.0032)	−0.0017 (0.0041)	−0.0017 (0.0041)	−0.0008 (0.005)	0.0015 (0.0044)
urban		0.0031*** (0.0011)	0.0013 (0.00084)	0.0004 (0.0008)	0.0004 (0.0008)	−0.00016 (0.0014)	0.00105 (0.0011)
roaddens		−0.000211** (9.01e−05)					
popdens			−2.91e−05 (1.96e−05)				
govsize				−0.0002 (0.0006)	−0.0002 (0.0006)	−0.0002 (0.0007)	−0.0023*** (0.0009)
oecd						−0.0236 (0.0361)	
eu							0.0586** (0.0257)
Constant, time dummies							
Observations	149	63	146	96	96	96	96

Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.9: Pooled robust regressions – freeness of trade – Lessmann

Variable	(1) PoolP1 covw	(2) PoolP2 covw	(3) PoolP3 covw	(4) PoolP4 covw	(5) PoolP5 covw	(6) PoolP6 covw	(7) PoolP7 covw	(8) PoolP8 covw
phiness	0.0782*** (0.0275)	0.0855*** (0.0289)	0.134*** (0.0365)	0.0933*** (0.0239)	0.0453 (0.0365)	0.110** (0.0510)	0.0934*** (0.0320)	0.107*** (0.0350)
loggdppc	0.259*** (0.0548)	0.318*** (0.0694)	0.179 (0.359)	0.197 (0.133)	0.135 (0.0963)	0.302 (0.526)	0.330*** (0.0683)	0.371*** (0.0770)
loggdppc2	-0.0187*** (0.00340)	-0.0217*** (0.00400)	-0.0149 (0.0194)	-0.0152** (0.0072)	-0.0127** (0.0052)	-0.0218 (0.0295)	-0.0217*** (0.0042)	-0.0249*** (0.0044)
urban		-0.0009 (0.0006)	0.0008 (0.0017)	-0.0004 (0.0007)	-0.0006 (0.0006)	-0.0005 (0.0011)	-0.0012* (0.0007)	-0.0009 (0.0008)
roaddens			-0.0002 (0.0003)					
popdens				-8.99e-05** (4.22e-05)				
agri					-0.0053*** (0.002)			
govsize						-0.0002 (0.0071)		
oecd							-0.0380 (0.0535)	
eu								-0.0497** (0.0216)
Constant, logunits, logarea, logare/logunits, time dummies								
Observations	194	194	73	194	189	113	194	194
r2 w	0.774	0.771	0.759	0.786	0.790	0.732	0.759	0.789
r2 rho	0.433	0.439	0.466	0.458	0.446	0.448	0.446	0.449

Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Robust regression with outlier detection

Table A.10: FE – freeness of trade – Lessmann

Variable	(1) FEP1 covw	(2) FEP2 covw	(3) FEP3 covw	(4) FEP4 covw	(5) FEP5 covw	(6) FEP6 covw	(7) FEP7 covw	(8) FEP8 covw	(9) FEP9 covw
phiness	0.0690 (0.0692)	0.0594 (0.0602)	0.0103 (0.0918)	0.0103 (0.0918)	0.0585 (0.0594)	0.0552 (0.0600)	0.0390 (0.0543)	0.0390 (0.0543)	0.0390 (0.0543)
loggdppc	0.187 (0.160)	0.259* (0.133)	-0.0159 (0.308)	-0.0159 (0.308)	0.395* (0.235)	0.215 (0.159)	0.493** (0.240)	0.493** (0.240)	0.493** (0.240)
loggdppc2	-0.0098 (0.0093)	-0.0143 (0.0088)	0.0061 (0.0234)	0.0061 (0.0234)	-0.0232 (0.0156)	-0.0119 (0.0097)	-0.0270* (0.0153)	-0.0270* (0.0153)	-0.0270* (0.0153)
urban		-0.0057* (0.0029)	-0.0034 (0.0036)	-0.0034 (0.0036)	-0.0055* (0.0029)	-0.0074** (0.0036)	-0.0106** (0.0053)	-0.0106** (0.0053)	-0.0106** (0.0053)
roaddens			-0.00015 (0.00027)	-0.00015 (0.00027)					
popdens					-0.0002 (0.0002)				
agri						-0.0022 (0.004)			
govsize							-0.0035* (0.0018)	-0.0035* (0.0018)	-0.0035* (0.0018)
Constant, logunits, logarea, logarea/logunits, two way FE									
Observations	194	194	73	73	194	189	113	113	113
R^2	0.431	0.472	0.356	0.356	0.484	0.483	0.563	0.563	0.563
F	4.921	5.140	15242	15242	7.191	5.250	4.878	4.878	4.878

Robust standard errors in parentheses; ** $p < 0.01$, * $p < 0.05$, * $p < 0.1$

Table A.11: Panel IV – freeness of trade – Lessmann

Variable	(1) PanIVP1 covw	(2) PanIVP2 covw	(3) PanIVP3 covw	(4) PanIVP4 covw	(5) PanIVP5 covw	(6) PanIVP6 covw	(7) PanIVP7 covw	(8) PanIVP8 covw	(9) PanIVP9 covw
phiness	0.171 (0.130)	0.0624 (0.126)	0.0369 (0.234)	0.0369 (0.234)	0.0454 (0.125)	0.0499 (0.130)	0.167* (0.0978)	0.167* (0.0978)	0.167* (0.0978)
loggdppc	0.226** (0.107)	0.260** (0.102)	0.0457 (0.612)	0.0457 (0.612)	0.391*** (0.128)	0.213* (0.121)	0.598*** (0.184)	0.598*** (0.184)	0.598*** (0.184)
loggdppc2	-0.0143* (0.0084)	-0.0144* (0.008)	0.00158 (0.0445)	0.00158 (0.0445)	-0.0227** (0.0094)	-0.0117 (0.0086)	-0.0348*** (0.0127)	-0.0348*** (0.0127)	-0.0348*** (0.0127)
urban		-0.0057*** (0.0018)	-0.0035 (0.005)	-0.0035 (0.005)	-0.0055*** (0.0018)	-0.0074*** (0.0021)	-0.0091** (0.0045)	-0.0091** (0.0045)	-0.0091** (0.0045)
roaddens			-0.00016 (0.0007)	-0.00016 (0.0007)					
popdens					-0.0002* (0.0001)				
agri						-0.0022 (0.0025)			
govsize							-0.003 (0.00192)	-0.003 (0.0019)	-0.003 (0.0019)
Constant, logarea, logarea/logunits, FE, time dummies									
Observations	194	194	73	73	194	189	113	113	113
r^2 w	0.405	0.472	0.350	0.350	0.483	0.483	0.506	0.506	0.506
F	31.27	33.86	48.76	48.76	32.51	32.63	29.38	28.52	29.20

Standard errors in parentheses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.12: System GMM – freeness of trade – Lessmann

Variable	(1) SysGMMP1 covw	(2) SysGMMP2 covw	(3) SysGMMP3 covw	(4) SysGMMP4 covw	(5) SysGMMP5 covw	(6) SysGMMP6 covw	(7) SysGMMP7 covw	(8) SysGMMP8 covw
L.covw	0.984*** (0.0841)	0.903*** (0.0421)	1.458*** (0.160)	1.374*** (0.116)	0.907*** (0.0449)	0.955*** (0.0615)	0.972*** (0.0599)	0.979*** (0.0569)
phiness	−0.0087 (0.0186)	−0.0011 (0.0192)	−0.0133 (0.0177)	−0.0044 (0.0147)	−0.00546 (0.0163)	−0.007 (0.0157)	−0.0052 (0.0165)	−0.0237 (0.0167)
loggdppc	0.0425 (0.0507)	0.0422 (0.0527)	0.0249 (0.0587)	−0.0623 (0.0381)	−0.0465 (0.0583)	0.0379 (0.0679)	0.0369 (0.0708)	−0.0672 (0.0757)
loggdppc2	−0.003 (0.0034)	−0.0036 (0.0032)	−0.0014 (0.0033)	0.0031 (0.0023)	0.0009 (0.0034)	−0.0033 (0.004)	−0.0031 (0.0039)	0.0026 (0.0043)
urban		0.0004 (0.00066)	−0.00026 (0.0007)	0.00035 (0.0003)	−8.60e−05 (0.0007)	0.0001 (0.001)	9.25e−05 (0.001)	0.0011 (0.0011)
L2.covw			−0.429*** (0.152)	−0.411*** (0.122)				
roaddens			−9.50e−06 (5.15e−05)					
popdens				−1.05e−05 (1.15e−05)				
agri					−0.0039** (0.0016)			
govsize						0.0002 (0.0007)	0.00025 (0.0006)	−0.0024** (0.0011)
oecd							−0.00293 (0.0271)	
eu								0.0776*** (0.0286)
Constant, logunits, logarea, logarea/logunits, time dummies								
Observations	144	144	50	91	141	94	94	94
Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$								

Table A.13: IV regression with lagged GDP

Variable	(1) IVLE1 covw	(2) IVLE2 covw	(3) IVLE3 covw	(4) IVLE4 covw	(5) IVLE5 covw	(6) IVLE9 covw	(7) IVLE10 covw
tradegdp	0.272*** (0.0446)	0.305*** (0.0482)	0.407*** (0.0814)	0.303*** (0.0452)	0.274*** (0.0485)	0.327*** (0.0553)	0.320*** (0.0524)
loggdppc	0.161** (0.0672)	0.297*** (0.0842)	0.307* (0.170)	0.200** (0.0988)	0.143 (0.109)	0.436*** (0.122)	0.517*** (0.125)
loggdppc2	−0.0130*** (0.0041)	−0.0198*** (0.0049)	−0.0199* (0.0103)	−0.0142** (0.0057)	−0.0121** (0.006)	−0.0277*** (0.0072)	−0.0297*** (0.0071)
urban		−0.0026*** (0.00084)	−0.0027** (0.0013)	−0.0022** (0.00091)	−0.0029*** (0.00097)	−0.0038*** (0.001)	−0.0048*** (0.0011)
roaddens			−6.78e−05 (0.0002)				
popdensity				−0.0001*** (3.04e−05)			
agri					−0.0061*** (0.0022)		
govsize						−0.00289* (0.00138)	−0.0023* (0.0012)
oecd							−0.126*** (0.0332)
Constant, logunits, logarea, logarea/logunits, time dummies							
Observations	194	194	73	194	189	113	113
R^2	0.646	0.664	0.651	0.680	0.675	0.692	0.723
F	48.49	37.69	12.24	35.65	35.07	22.99	23.04
Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$							