Migration and Regional Trade Agreement: a (new) Gravity Estimation

Abstract

This paper investigates the role of Regional Trade Agreements (RTAs) on bilateral international migration. Building on the gravity model for migration by Anderson (2011), our econometric strategy controls for the multilateral resistance to migration (Bertoli and Fernandez-Huertas Moraga, 2013) and solves the zero migration flows problem by using a censored quantile regression approach. Further, the endogeneity problem of RTAs in migration settlement is addressed by using IV censored quantile regression (Chernozhukov and Hansen 2008). Our results suggest that the presence of a RTA stimulates the migration stocks among member countries. The pro-migration effect of RTAs is magnified if the agreement includes also provisions easing bureaucratic procedures for visa and asylum among member countries. Finally, we find a asymmetric effect of RTAs across the quantiles of the distribution of migration settlements.

JEL classification: C13, C23, F22, F13

Keywords: Migration, Gravity Equations, Censored Quantile Regression

1 Introduction

Over the past 40 years, the gravity equation has been the workhorse for trade economists studying the effect of trade liberalization on bilateral trade flows. As recently highlighted by Head and Mayer (2014), given its high flexibility, the gravity micro-foundation can be applied to a wide range of bilateral flows and interactions. Anderson (2011) extended the micro foundation of the gravity model to both foreign direct investment (FDI) and migration flows. This paper focuses on the latter domain and uses gravity model to explain international migration settlements across countries as a function of migration costs. In particular we focus on the role played by Regional Trade Agreements (RTAs) as source of information about destination countries for potential migrants in origin countries. This is expected to reduce migration cost. The choice of the destination country implies in comparisons of several potential destinations in terms of the expected utility level they generate (expected income)[1]. Such comparison requires information on income, condition of living and adaptation costs at each potential destination, but such an information is often hard to find (i.e. *information cost of migration*). However, RTAs, by raising the awareness of the partner countries reduce the information cost of migrating to RTA's member countries.[2] Beyond the information channel, RTAs may affect international migration by simply stimulating trade flows. Indeed, in a standard factor content model of trade, RTAs boost trade flows and thus reduce wage inequalities among member countries, which in turn reduces the incentive for international migration. If the latter effect prevails, then we should observe a negative relation between RTAs and bilateral migration flows. However, previous evidence (i.e., Orefice 2015) shows that, all other determinants constant, RTAs and bilateral migration flows are positively related, suggesting that RTAs drive migration choice towards member countries.[3]

[1]The Random Utility Model (RUM) for migration predicts that the potential migrant will migrate in the country that maximises his/her expected utility level.

[2]Indeed, signing a RTA leads to improved diplomatic relations and greater familiarity among signatory countries.

[3]An increasing number of recent RTAs also include visa and asylum provisions aimed at easing bilateral bureaucratic cost of migration (see Horn, Mavroidis and Sapir 2010 and Orefice 2015 for detailed discussion on this point).

This paper adds to Orefice (2015) in two different ways. First, we apply a fresh econometric technique to address recent problems observed in the Pseudo Poisson Maximum Likelihood (PPML) estimator. Indeed, PPML, being a non linear estimator, over-weights large bilateral migration flows (Head and Mayer 2014). This is a crucial shortcoming in gravity for migration because very big bilateral migration flows usually involve the same country pairs across years (i.e., Mexico-USA, Morocco-Spain, Poland-Germany). Such specific migration cases are due to historical and geographic factors and do not depend on time varying bilateral migration costs (i.e. the information cost); for this reason, we expect that the PPML estimator provides biased coefficient on the effect of RTAs on migration settlement. Moreover, Figueiredo, Lima and Schaur (2014) show that the PPML relies on identification conditions that may not hold in practice, leading to a significant bias in the estimates of the gravity equation. Our second innovation is to enlarge the sample of destination countries covered in the estimation of the gravity model. Thank to the recent World Bank data on bilateral migration stocks, we are able to cover developing (destination) countries and thus include south-south migration flows in our study. Hence, we are able to investigate whether RTAs affect differently the pattern of migration towards developing countries.[4] This is an important feature given the recent tendency of developing countries in signing RTAs.[5]

The theoretical micro-foundation of the gravity equation for trade provided by Anderson and Van Wincoop (2003) highlighted the importance of controlling for the multilateral price resistance term (MRT) in gravity equation estimation (see Head and Mayer 2014 and Baltagi, Egger and Pfaffermayr 2014 for an exhaustive survey of the literature on this point). The MRT term can be captured by country specific fixed effects in cross sectional samples, or by country-by-year fixed effects in case of panel data samples. In a parallel way, the micro-foundation of gravity equation for migration proposed by Anderson (2011)

[4] We thank an anonymous referee for suggesting us this extension.

[5] The list of early announcement - under negotiation - RTAs, provided by the WTO RTAs information system, includes a lot of agreements involving developing countries, such as Canada - Dominican Republic launched on June 7th, 2007, EFTA - Indonesia launched on January 31st, 2011, EU - Thailand launched on March 6th, 2013, Ukraine - Singapore launched on May 8th, 2007, and many others available at: http://rtais.wto.org/UI/PublicEARTAList.aspx.

highlights the importance of migration costs and inward/outward migration frictions in predicting migration flows and their settlements across destination countries.[6] The importance of MRT to migration has also been highlighted by Bertoli and Fernández-Huertas Moraga (2013) who show that the multilateral resistance to migration is properly captured by country-by-year fixed effects in a standard random utility maximization (RUM) model for migration where the attractiveness of a destination country does not vary across origin countries (see also Beine, Bertoli and Fernández-Huertas Moraga 2014 on this point). Inward migration frictions – *Inward migration resistance term* - may be imagined as immigration policy restrictiveness; while outward migration frictions – *outward migration resistance term* – can be considered as origin country specific factors deterring emigration.[7] For instance, the emigration frictions from an origin country (*outward migration resistance term*) depend on its remoteness with respect to the rest of the world, and/or on its travel connections with the potential destinations which may vary over time (train or navy routes); also the average income in the origin country might represent a friction in emigrating when it is too low to overcome the migration travel costs. Failing to control for such migration frictions produces an omitted variable problem and consequently leads to biased estimators (as in the case of gravity for trade). Thus, in our empirical gravity for migration we properly capture inward/outward migration frictions by including country-by-year fixed effects.

Although the inclusion of country (or country-by-year) fixed effects addresses the omitted variable problem, an endogeneity issue may still arise due to the presence of reverse causality. Ghosh and Yamarik (2004) and Baier and Bergstrand (2007) highlighted the endogeneity of RTAs in the context of gravity for trade.[8] However, the reverse causality

[6]In Anderson (2011) the decision to migrate as an individual discrete choice model conditioned to country pair specific migration cost, and aggregated migration flows are determined by the relative size of origin and destination country, migration cost and inward/outward migration frictions (see eq. 20 in Anderson, 2011)

[7]Inward and outward migration resistance terms highlighted by Anderson (2011) mirror, to some extent, the more traditional push and pull factors setting used to explain bilateral migration flows (Hatton, 2005; Mayda, 2010; Grogger and Hanson, 2011).

[8]Indeed, RTAs formation may be affected by trade flows (to secure current bilateral flows), leading to endogeneity of RTAs in the estimation of gravity equations.

argument might also apply in the context of gravity for migration when a RTA is signed as a consequence of high migration pressure.[9] Another econometric issue extensively discussed in the trade literature (but scarcely in the context of migration) is the zero migration flows problem. If the dependent variable has a large share of zeros (as in the case of bilateral migration) and the data exibits heteroskedasticity, then the standard OLS estimator will be biased (see section 2.2 for more details).

This paper addresses the previous identification issues by using recent developments in instrumental variable censored quantile regressions proposed by Chernozhukov and Hansen (2008), Chernozhukov, Fernandez-Val and Kowalski (2011), Kowalski (2008), and Kowalski (2013). We use data from the World Bank for the 1960-2010 period and estimate the effect of RTAs on the entire distribution of bilateral migration stocks (see section 3.2 for a detailed description of the data and discussion on the use of migration stocks). By using a quantile approach, we also test any potential asymmetry in the RTAs' effect on migration (the idea is that RTAs may have a different effect along the distribution of migration stocks). Finally, we study whether the inclusion of visa and asylum provisions in the RTA changes its effects on bilateral migration settlement.

We find that, failing to control for multilateral resistance to migration produces very unstable RTAs coefficients (in particular for OLS and PPML), and introduces a downward bias in the Censored Quantile Models - CQ Model (our preferred specification). After the inclusion of the country-by-year specific fixed effects, the CQ model shows positive and significant RTA coefficient. According to our preferred specification, the presence of a common RTA stimulates bilateral migration by 8%. After controlling for endogeneity (IV on censored quantile regression model), the effect of RTAs on migration stocks gets bigger, suggesting that RTA dummy is indeed endogenous with respect to migration stocks. Finally we show that the RTAs effect is increasing over the moments of the distribution of migration stocks: the signature of RTAs has a null effect on migration stocks at the 25^{th}

[9] In a standard Heckscher-Ohlin framework, trade in goods is supposed to deter international migration flows by equalizing factor prices. So in this case, government may decide to boost trade in goods to avoid migration flows.

percentile of the distribution, but it has an increasing positive effect over the remaining right side of the distribution.

In the next section we describe in details the three possible shortcomings in estimating gravity equation for migration. Section 3 presents the details of our preferred estimator. In section 4 we present our results. Last section concludes.

2 Gravity Model for Migration and Bilateral Migration Costs

In this paper we estimate a robust gravity model for migration focusing on the effect of Regional Trade Agreements (RTAs) on bilateral international migration.[10] The idea is that RTAs increase the awareness of the partner country among the potential migrants: the signature of a RTA implies improved diplomatic relations among signatory countries, and thus improves the information for all potential migrants in the destination country. This reduces bilateral migration costs and, all other determinants being constant, favors migration among signatory countries. Indeed, the signature of RTAs implies more frequent diplomatic relations and a different perception of the (new) member countries as "friend" countries. Moreover, the signature of RTAs implies a wide media coverage of the issue (newspapers and TV networks report the news frequently) implying more information on the economic environment of the member countries. Unfortunately we do not have data to directly test this channel, but, after controlling for the presence of migration specific provisions in RTAs, this channel remains a plausible explanation for the positive effect of RTA on bilateral migration stocks.

In what follows, we explain the role of bilateral (time varying) migration cost (RTAs) in affecting the pattern of migration. To this end we rely on Anderson (2011) for the theoretical foundation of the gravity equation for migration. Bilateral migration costs are dyadic factors deterring migration flows, in particular, such costs may vary across years

[10] For a recent and updated practitioners' guide to gravity model for international migration see Beine, Bertoli and Fernández-Huertas Moraga (2015).

or being time-invariant. Time-invariant migration costs are country pair specific factors affecting migration, in general one may imagine such cost as geography driven. Linguistic barriers, colonial links, distance (as a proxy for travel cost) and cultural proximity belong to this kind of costs. Time variant migration costs refer to country pair factors that may change across years, information cost and policy barriers belong to this kind of costs. Recently, Bertoli and Fernandez-Huertas Moraga (2012) used bilateral visa policies as a proxy for bilateral policy measures.[11] Here we use data on the content of RTAs (namely on provisions easing the bureaucratic procedure to obtain a visa) to approximate the policy related cost of migration.

Information cost is a further, important, determinant of migration. The potential migrant chooses his/her destination on the basis of the information he/she has on the destination country. In the literature, such an information cost has been often approximated by past stock of migrants (network effect) - measured as the stock of immigrants in the starting year. However such a proxy, being time-invariant, does not properly control for time varying information cost of migration (information about a destination country is likely to change over time). We consider the presence of a common RTA as providing additional information about the destination country to potential migrants in the origin country.

Considering the previous framework, and simply relying on Anderson's (2011) gravity equation, bilateral migration settlements can be expressed as:

$$M_{ij} = \frac{N_i N_j}{N}\left(\frac{\delta_{ij}}{\Gamma_i \Omega_j}\right)^{1-\theta}, \tag{1}$$

where

$$\Gamma_i \equiv \sum_j \left(\phi_j/\delta_{ij}^{1-\theta}\right)\Omega_j^{\theta-1}, \tag{2}$$

$$\Omega_j \equiv \sum_i \left(\phi_i/\delta_{ij}^{1-\theta}\right)\Gamma_i^{\theta-1}. \tag{3}$$

[11] Authors focus on the Spanish immigration case.

M_{ij} represents the number of migrants living in country i coming from j; N_i and N_j indicate respectively population size in destination and origin country, while N is the total population size. The Γ_i is the appropriate 'average' portion of migration costs borne by country i to all destinations, outward multilateral resistance, and Ω_j is the average portion of migration costs borne by j from all sources, inward multilateral resistance, with $\phi_i = N_i/N$, $\phi_j = N_j/N$. δ_{ij} represents bilateral migration costs.

The equation (1) is exactly analogous to the Anderson and van Wincoop (2003) gravity model for trade, while equations (2) and (3) respectively trace inward and outward multilateral resistance equations for trade (MRT), but applied to migration. Thus, the MRT are not observable but can be inferred along with δ_{ij}. The equation (1) can be split in two parts: $N_i N_j/N$ is the frictionless share of migrants in country j and $(\delta_{ij}/\Gamma_i\Omega_j)^{1-\theta}$ is the effect of migration frictions.

Following the literature, we extend equation (1) to a time variant setting by including a t subscript and allowing both inward and outward resistance term to migration to vary over time. For this reason in our empirical exercise we control for inward/outward resistance term by including country-by-year fixed effects. It must be noticed that the inclusion of country-by-year fixed effects properly controls for the multilateral resistance to migration only under the assumption that the attractiveness of the destination country does not vary across origin countries (see Bertoli and Fernandez-Huertas Moraga 2013, and Beine, Bertoli and Fernandez-Huertas Moraga 2014 for more details on this point). However, the inclusion of country-by-year fixed effects to control for the multilateral resistance term to migration has already been used in the recent literature on the determinants of bilateral migration (among others Beine and Parsons 2012; Ortega and Peri 2013) and represents a compelling way to control for the multilateral resistance to migration.

Also bilateral migration costs can be imagined as varying over time. In particular we imagine the following exponential function for migration cost:

$$\delta_{ijt} = FC_{ij}^{\beta_1} VC_{ijt}^{\beta_2}, \tag{4}$$

where FC_{ij} is the time invariant component of bilateral migration cost such as cultural diversity between origin and destination country (based on past migration settlement, common language or past colonial relationship), and geographical cost of migration (i.e., distance and common border); VC_{ijt} is the time variant cost of migration based on the information cost and β_1 and β_2 are parameters.

In what follows, we capture time invariant migration costs (both geography and culture diversity driven) by using bilateral distance (travel cost), common language, common border, past colonial link dummies and past stock of migrants. Time variant migration cost are approximated by RTA_{ijt}, our crucial variable, which is a dummy variable equal to one if country i and j share a common RTA at time t. As a proxy for the bilateral policy related (and time-variant) migration cost we use a dummy variable equal to one if the RTA contains a visa-and-asylum provision. The visa and asylum provision is meant to reduce the bureaucratic cost for obtaining a visa among signatory countries, so its inclusion in the RTA is a further incentive for the potential migrant to choose his/her destination among signatory countries.[12] In our specifications we also include two other control variables: a WTO dummy, being equal to one if country i and j are both part of the WTO, which isolates the "preferential" nature of the RTA; and the difference in per capita GDP between origin and destination countries in order to capture the effect of the relative attractiveness of the destination.[13]

Most of the existing studies on the determinants of international pattern of migration use the flows of migration (IMD database) provided by the OECD (the use of migration flows is coherent with Anderson 2011). This allows to have annual data on migration flows from origin to destination country, but it also restrains the analysis to a short time period and, more important, on OECD destination countries only. This would be an important

[12]The presence of a Visa-asylum provision in the RTA is based on a recent mapping of 96 RTAs provided by the WTO. This dataset is not exhaustive of all RTAs, this is why our specifications using visa-and-asylum dummy have a reduced number of observations. Such database is available at: http://www.wto.org/english/res_e/publications_e/wtr11_dataset_e.htm.

[13]Controlling for the relative attractiveness of countries is important because, as highlighted by Bertoli and Fernández-Huertas Moraga (2013), the multilateral resistance to migration is properly captured by country-by-year fixed effects if the attractiveness of a destination country does not vary across origin countries.

limitation for our purpose since it prevents us from covering south-south migration flows, for which the information channel is plausibly more important. For these reasons, we decided to rely on bilateral stock data on migration provided by the World Bank. This reduces the number of observations across time, but guarantees deeper historical perspective (our database starts in 1960) and a wider set of destination countries (with respect to the OECD data). Some previous studies use differences in stocks as a proxy for (long run) flows, however in this way the dependent variable may assume negative values (due to returning migration, deaths and naturalization) creating problems in terms of estimation and intuition (migration in the Anderson 2011 model is strictly positive). Thus, in this paper we follow Grogger and Hanson (2011) and Llull (2014) and simply use the stock of migrants in country i coming from j at time t. [14]

In the next section we focus on the potential endogeneity problem caused by both omitted variables and reverse causality and then we move to the issue caused by the presence of migration stocks equal to zero.

2.1 Endogeneity Problem: Omitted Variable and Reverse Causality

The first crucial econometric issue in estimating gravity model for trade is the endogeneity related to the RTA variable due to omitted variables and reversal causality problems. The two previous problems hold also in a migration based gravity model.

The omitted variable problem comes directly from the theoretical micro-foundation of gravity for migration provided by Anderson (2011), where bilateral migration flows (and thus stocks) depend on destination and origin country specific migration frictions (outward and inward migration resistance term).[15] Such migration frictions are unknown to the econometrician and thus likely to be omitted in a gravity estimation model; ignoring

[14]As a robustness check, in the appendix section A1 we do the same exercise by using migration flow data obtained from the OECD IMD database.

[15]Bertoli and Fernández-Huertas Moraga (2013) define the multilateral resistance to migration as the contradictory effect that the attractiveness of alternative destinations exerts on the determinants of bilateral migration flows.

outward and inward migration resistance terms generates bias in the estimated coefficients. In this paper we account for outward and inward migration resistance term by including country-by-year fixed effects. By comparing a model including only country fixed effect with a model including country-by-year fixed effects we may also provide an estimation of the bias caused when one omits outward and inward migration resistance term in gravity for migration.

The reverse causality problem is related to the possibility that RTAs are signed in response to migration pressure. We address this problem by using an instrumental variable approach, which is fully explained in section 3. Our instrumental variable is based on the idea of a domino effect in RTAs formation highlighted by Baldwin and Jaimovich (2012) – and already used in Orefice (2015) – who show that the probability that two countries join in a common RTA is positively affected by the number of RTAs that each potential partner has with the rest of the world (in order to avoid trade diversion effect). Following this idea, we use the total number of RTAs signed by origin and destination country (minus one if they have a RTA in common) to instrument the common RTA dummy. The number of RTAs signed by each country with the rest of the world can be thought as exogenous with respect to bilateral migration flows since having a RTA with a third country does not affect the inflow of migrants from the partner country j. The exclusion restriction for our instrument is related to the fact that having an existing RTA with a third country z does not affect the bilateral specific migration flows ij. We also test the validity of the exclusion restriction in what follows.

2.2 The Zero Migration Problem

The presence of a large amount of observations equal to zero in trade flows has been highlighted by Helpman, Melitz and Rubinstein (2008); Santos Silva and Tenreyro (2006); Head and Mayer (2014) and Anderson (2011). Indeed, Anderson (2011) argues that many potential bilateral trade flows are not active. The data presented to the analyst may record a zero that is a true zero or it may reflect shipments that fall below a threshold above zero.

Helpman, Melitz and Rubinstein (2008) show that country pairs with zero trade account for about half of the observations. The presence of a high share of observations equal to zero in the dependent variable calls for two needs: (a) considering observations equal to zero since they are important source of information; (b) robust estimation techniques in the presence of zeros. The problem of zeros is crucial also in a migration setting where plausibly for a consistent share of country pairs a substantial migration stock does not exist. This is indeed what we observe in our sample of migration stocks; the continuous line in Figure 2 shows the empirical distribution of the logarithm of bilateral migration stocks and the big mass at zero.

Moreover, OLS estimation of the log-linear model of migration might suffer from heteroskedasticity in the error term (see Santos Silva and Tenreyro, 2006). Among the several estimators proposed in the literature to solve both heteroskedasticity of error terms and the presence of zeros, there is a fair understanding in the trade literature in considering a Poisson Pseudo Maximum Likelihood (PPML) as the most appropriate estimator.[16]

However PPML, being a non-linear estimation in levels, tends to over-weight large bilateral flows, which is a relevant issue in the case of migration stocks. Indeed, the largest migration stocks are due to some specific country pairs (destination-origin), for example: (1) USA - Mexico; (2) Germany - Poland; (3) Spain - Morocco. So a PPML will tend to over estimate the population of migrants as shown in Figure 2. To avoid this problem we propose here a further estimator which produces unbiased and consistent estimates in presence of high share of zero stocks and does not over weight large bilateral stocks. Additionally, the proposed estimator correctly accounts for endogeneity, heteroskedasticity and allow us to investigate the effect of RTA on the entire distribution of migration.

[16] See Head and Mayer (2014) for a survey on the estimation techniques proposed in trade literature to solve the zero trade flow problem.

3 A New Gravity Estimation

3.1 The Econometric Model

In specific notation, we consider the following exponential model studied by Santos Silva and Tenreyro (2006)

$$M_{ij} = \exp(x_{ij}\beta)\,\eta_{ij}, \tag{5}$$

where, in this paper, M_{ij} is the migration stock from i to j, x_{ij} represents the explanatory variables and the multilateral resistance terms, β is a vector of parameters and, η_{ij} is a non-negative random variable. We can linearize the model by taking logarithms of both sides of the equation to obtain

$$\ln M_{ij} = x_{ij}\beta + \ln \eta_{ij}, \tag{6}$$

where $\ln M_{ij}$ is now defined on the real line $\mathbb{R}$.

Heteroskedasticity can be included in this model by assuming that $\eta_{ij} = \exp[(x_{ij}\gamma)\,\varepsilon_{ij}]$, where ε_{ij} is i.i.d.. In this case, the above model becomes

$$\ln M_{ij} = x_{ij}\beta + (x_{ij}\gamma)\,\varepsilon_{ij}. \tag{7}$$

This is a location-scale model, in which the covariates x_{ij} affect not only the location (mean) of the conditional distribution of $\ln M_{ij}$, but also its scale and quantiles through $(x_{ij}\gamma)$.

Santos Silva and Tenreyro (2006) argued that if $E(\eta_{ij}|x) = 1$ (which implies that model (5) is identified) and $\gamma \neq 0$ (meaning that there exists heteroskedasticity), then the OLS estimator applied to the log-linear model (7) is severely biased. More recently, Figueiredo, Lima and Schaur (2014) showed that if we do the other way around by first assuming that model (7) is identified (which implies that $E(\varepsilon_{ij}|x) = 0$), then the PPML estimator considered by Santos Silva and Tenreyro (2006) will be severely biased. In other words, it is not possible to say which estimator is biased because we do not observe the identifying condition of each estimator. More importantly, by the Jensen's inequality $E(\ln \eta_i | x_i) \neq \ln [E(\eta_i | x_i)]$ meaning that identification of the exponential model (5) does not lead to identification of the log-linear model (7) and vice-versa.

To address the above problem, Figueiredo, Lima and Schaur (2014) proposed using quantile regression to identify both models (5) and (7). Their idea relies on the fact that, unlike the mean function, the quantile function is invariant to monotone transformations. In other words, if $h(\cdot)$ is a nondecreasing function on $\mathbb{R}$, then for any random variable Y, $Q_\tau(h(Y)) = h(Q_\tau(Y))$,where $Q_\tau(\cdot)$ is the $\tau - th$ quantile function. Based on this property, they show that identification of the exponential model leads to identification of the log-linear model and vice-versa without assuming any knowledge about the distribution function of η_{ij}. Because the quantile approach identifies both models, we can focus our estimation on the log-linear model to avoid the problem of overweighting large bilateral migration stocks. However, as in the trade literature, the main problem related to the log-linear model (6) is a large amount of zero migration stocks (see Orefice 2015, Ramos and Suriñach 2013, and among others). For these observations, taking the log of zero will automatically lead the computer to drop them. To address this issue, we consider the "Tobit for lognormal model" as explained in Cameron and Trivedi (2009). In this framework, a nonnegative latent variable (M_{ij}) is defined by an exponential model (5), but we only observe

$$M_{ij}^* = \begin{cases} M_{ij}, & \text{if } \ln M_{ij} > \kappa \\ 0, & \text{if } \ln M_{ij} \leq \kappa \end{cases}$$

Therefore, in this model "zeros" result from statistical rounding. In other words, one rounds M_{ij} down to zero whenever $\ln M_{ij} \leq \kappa$ (or $M_{ij} \leq \exp(\kappa)$). Statistical rounding is common in the literature of gravity equation as mentioned by Head and Mayer (2013) and Head, Mayer and Ries (2010).

Cameron and Trivedi (2009, page 545) suggest that the censoring point, κ, should be set to a value equal to or less than the *minimum uncensored value* of $\ln M^*_{ij}$. In our database, we use the one-thousand scale and we found the minimum uncensored value of migration stocks to be equal to 1, which leads us to set $\kappa = 0$.[17] Thus, in what follows we will set the estimated log-linear model as

$$\ln z_{ij} = \max\left(0, \ln M_{ij}\right) \tag{8}$$

which corresponds to the censored regression model considered by Dutt, Mihov and Zandt (2011) in the trade literature.

In this paper, estimation and inference is conducted by using a generalization of the Tobit model developed by Powell (1984, 1986) which, unlike the Tobit model, identifies the parameters of interest without imposing normalility and homoskedasticity. Finally, as shown in the subsection 2.1, a successful estimator for the log linear gravity model for migration must also address the endogeneity issue. This is carried out by applying to the Powell estimator the IV moment condition approach developed by Chernozhukov and Hansen (2008) (see Appendix section for technical explanations on the econometric modelling).

In sum, the estimating strategy adopted by this paper is able to address all the identifying issues discussed in the previous sections, which includes the presence of a large number of observations close and equal to zero, overweighting of large migration stocks, endogeneity of the RTA, and heteroskedasticity. Moreover, the proposed estimator will allow us to investigate the effect of RTA not only at the mean, but also at the entire distribution of bilateral migration settlements. In what follows, we present the data used in our empirical exercise as well as our main results.

[17] This corresponds to the minimum uncensored value of $\ln M^*_{ij}$ and therefore satisfies the rule proposed by Cameron and Trivedi (2009)

3.2 Data

We use data from World Bank Bilateral Migration Matrix 2010 (Özden et al, 2011). We have data for 200 countries (see Table 1) for a long time period starting in 1960 and ending in 2010 and it provides information on bilateral migration stocks for every 10 years: 1960, 1970, 1980, 1990, 2000 and 2010. The number of country-pairs and years covered by the World Bank Bilateral Migration Matrix suggests high (potential) number of observations. When we include the stock of immigrants in 1960 as a control variable (to control for the initial stock of migrants), we restrain the analysis to those country-pairs having non-missing stocks in 1960 over 5 periods (1970-2010), i.e. 88,273 potential observations. However, the dataset is also unbalanced and by excluding missing observations we end up with 80,345 observations. As in Ramos and Suriãch (2013), we consider that the stock of migration can be interpreted as a representation of a long-term equilibrium, which is likely to be of higher quality than annual immigrant flow data. We are aware of the fact that the gravity model for migration as in Anderson (2011) predicts flows and not stocks. But bilateral migration flows are unfortunately available only for a sub-sample of OECD destination countries (OECD IMD dataset). This would prevent us from studying the different impact of RTAs for developing destination countries (see section 4.1), for which the information provided by RTAs is particularly useful. When a RTA is signed, it might have a positive effect on bilateral migration flows, but the resulting effect on stock strictly depends on the initial stock of migrants. Even by controlling for the very initial stock of migrants in 1960 (as we do in all our regressions)[18], it would be complicated to justify a constant proportional effect of RTAs on migration stocks. Indeed, RTAs are likely to have a positive impact in particular on those country pairs having already high stock of migrants. This is a further reason for running Censored Quantile Regression model, which allows us to investigate the effect of RTAs over the entire distribution of the size of bilateral stock, not only at its mean. However, for the interested reader we replicate our baseline estimations using flows instead of migration stocks. Results, reported in the appendix Table A.1 confirm our

[18] For this reason our estimation sample does not include the 1960.

results.

Table 1 about here

There is a large number of observations less than minimum uncensored value ($\exp(\kappa)$). We followed our strategy of assuming that $\ln(z_{ij}) = \max(\kappa, \ln(M_{ij}))$. As long as $\exp(\kappa) = 1$, this implies that $\ln(z_{ij})$ is set equal to 0 whenever the original observations are subject to censoring, or, whenever $M_{ij} < 1$. After this adjustment 42% of our sample are zeroes.

The control variables are from the standard sources. GDP and population for origin and destination countries are from World Bank's World Development Indicators (WDI). Data on distances, common border, language and colonial link are from CEPII.[19] WTO membership dummy simply indicates whether both origin and destination country belongs to WTO. The RTA dummy has been computed from the comprehensive list of active RTAs provided by the WTO - Regional Trade Agreement Information System.[20] This dummy equals to one if there is a RTA in force between origin and destination country in a given year and zero otherwise. In this respect, we do not differentiate RTAs by their degree of depth (i.e. free trade, economic integration or custom unions) because we simply care about the information provided by each RTA independently of its depth[21], but European Union being a deep custom union allowing for the free movement of people might bias our results. For this reason, in column (8) of our baseline Table 2, we also report results with a EU dummy variable (EU dummy is also included in Table 6, column 2).

As shown in Figure 1, the number of RTAs (and coherently the number of member countries) exponentially increased over the period 1960-2010. Namely, in the period 1975-1985 we had the first wave in RTAs' signature, from 40 RTAs in 1975 to 60 in 1985. But the real exponential increase in the number of RTAs started in the early nineties: from 68 in 1991 to 365 in 2010. Such exponential increase has been widely described in the World

[19]Available at http://www.cepii.fr/CEPII/fr/bdd_modele/presentation.asp?id=6

[20]Available here: `http://rtais.wto.org/UI/PublicMaintainRTAHome.aspx.`

[21]However, as a robustness check in Table (A.2) we split the RTA dummy into two components: (i) Custom Union dummy (CU) equal to one if the country pair belongs to a common Custom Union, and (ii) a PTA/FTA dummy equal to one if the country pairs belongs to a Preferential Trade Agreement or Free Trade Agreement. Results show that both CU and PTA/FTA have positive effect on bilateral migration. And, in line with the intuition, CU has a larger effect on migration than simple PTA/FTA.

Trade Report 2011. Our data also suggest that over the decade of RTA's signature, the bilateral stock of migrants almost doubled (unconditional average growth rate of 104%[22]) while the growth rate of bilateral migration one decade before the RTA signature is (on average) only 30%. This highlights how important is a signature of RTA for the bilateral migration stocks.

We also use a dummy variable to indicate whether a RTA includes legally enforceable visas and asylum provision.[23] This dummy is based on a mapping of only 96 RTAs provided by the WTO, that is the reason by which the number of observations dramatically reduces when we include such a dummy variable in the regression specification (from 80,345 to 16,126 observations).[24] Visa and asylum provision is aimed at reducing the administrative costs of migration by promoting the exchange of information among signatory countries and the drafting of legislation in the area of visas and asylum for migrants (see Horn et al. 2010 for more details).

4 Results

We estimate the gravity model (1) to examine the impact of trade agreements on the conditional mean and conditional quantiles of bilateral migration stocks. To this end, we first use pooled data with the robust estimator presented in the previous section, which properly accounts for zero migration values and heteroskedasticity (see Appendix section A2 for more details). Then we move to the second set of estimations where potential endogeneity of the RTA dummy is addressed by using a IV moment condition approach.

[22]This might appear a huge number, but it should be noticed that for a lot of country pairs the initial stock of immigrants counts only few hundreds of immigrants. This is likely to generate future huge growth rates when the stocks increases to a significant number of individuals.

[23]The legal enforceability of the provision is based on the language of the text used on the agreement and on the presence of a dispute settlement clause.

[24]See Orefice and Rocha(2014) for detailed discussion on the mapping of RTAs. Dataset available at: http://www.wto.org/english/res_e/publications_e/wtr11_dataset_e.htm.

4.1 Baseline Results

In our gravity specification, the RTA dummy is equal to one if origin and destination country share a RTA in the current year. Therefore, a positive coefficient on the RTA dummy means that RTA membership increases migration settlement among member countries. The behavior of such coefficient across quantiles allows us to analyse the effect of trade agreements on the different moments of the migration distribution. In this section our preferred specification is the simple censored quantile regression model (CS model) with no control for endogeneity (for more details see equations (A-3)-(A-4) in the appendix section) - we consider endogeneity in the next section.

We further compare this estimator with the PPML estimator to evaluate the bias due to the fact that PPML overweights large observations.[25] Then, we compare specifications including country and year fixed effects (columns 1-3 in Table 2) with specifications having country-by-year fixed effects (columns 4-8 in Table 2), in this way we have an idea of the potential bias due to the omission of inward/outward resistance term to migration. The OLS estimates are simply meant as a benchmark in which we will only consider the positive values of migration.

Table 2 shows our baseline results. According to our preferred specification, censored quantile regression estimation with country-by-year fixed effects (column 4), we find that signing a RTA stimulates the bilateral settlement of migrant by 8%. The rest of control variables have the expected sign. Distance negatively affects migration settlements, while past migrants stocks, common border and common language increase migration by reducing the bilateral (time invariant) cost of migration. Such results are robust to several robustness checks and estimators (see columns 1-8).

In columns 7-8 we include a further (potential) component of the time varying bilateral migration cost: a dummy being equal to one if the RTA (shared by origin and destination country) includes also a visa and asylum provision easing the bureaucratic cost of

[25]Figure 2 shows the observed and fitted values for three models (OLS, PPML and CS). By comparing the PPML fit with the observed migration distribution, the bias yielded by the PPML is evident.

migration. Results in columns (7)-(8) confirm our previous findings on the RTA dummy (stimulating bilateral migration) and show a positive coefficient on visa-and-asylum dummy after controlling for the EU dummy (one if origin and destination countries belong to the EU). In particular, results in column (8) suggest that the presence of a common RTA boosts bilateral migration stocks by 27%, such elasticity grows up to 37% if the RTA includes also a provision on visa and asylum, which is plausible if we think that RTAs with visa-and-asylum provision provide information on the destination country and also ease the bureaucratic procedure to migrate to member countries. When we include a visa-and-asylum dummy, the coefficient for the RTA dummy increases, but it must be noticed that the sample shrinks a lot since we are now using only those RTAs for which we have information on the contents (WTO dataset, see previous section). The increased coefficient for the RTA might raise a concern about the content of RTAs that do not have visa-and-asylum provision. Indeed, a high coefficient for the RTA dummy may be driven by the presence of other migration related provisions in the remaining RTAs.[26] In Figure A1 we show the content of the remaining RTAs (i.e. those RTAs that do not include visa and asylum provision). They contain mostly trade related provisions (tariff reduction in manufacturing and agriculture sectors) and do not contain other provisions for more labor mobility. Morever, its effect on migration is not different from the one found by Orefice (2015) using migration inflows in OECD countries.

All in all, we can conclude that the signature of a RTA, by reducing the bilateral (time varying) component of migration costs, increases migration stocks. Such effect is magnified by introducing in the RTA a provision on visa and asylum which reduces the bureaucratic cost of migration.

Table 2 about here

By comparing results in column (1) and (4) we have an idea of the potential bias due to the omission of proper controls for inward/outward resistance term to migration. After the inclusion country-by-year fixed effects, the coefficient on RTA increases by 35%

[26] We thanks an anonymous referee for suggesting this explanation

suggesting a huge bias in estimations that do not control for the multilateral resistance term to migration. Further, by comparing in turn columns (1) and (2) and then columns (4) and (5) we discover a bias in the PPML estimation. In column (2), PPML produces null coefficient on RTA dummy, implying zero effect of RTA on bilateral migration settlements; in column (5) PPML produces a large positive and significant coefficient for RTA. This suggests the weak robustness of PPML estimator in this setting. Conversely, censored quantile regression procedure gives a robust positive and significant coefficients on RTA dummy.

Figure 2 confirms the goodness of fit of the proposed estimator (censored quantile regression) as compared with the OLS and PPML. We estimated a non parametric kernel for the predicted values of bilateral migration stocks by using all the three estimators, and compare these with the observed distribution of migration stocks (all numbers are in the logarithmic scale).[27] If we consider the actual distribution of the migration observations, we notice that it is asymmetric with a mean value slightly above zero. We observe that the PPML produces a curve that is symmetric and shifted to the right of the actual distribution, reflecting the fact that PPML overweights large migration stocks that are present in the sample used in this paper. Figure 2 also shows that the curve obtained through OLS is symmetric and also biased to the right. The curve obtained by the proposed censored quantile regression not only recover the missing small observations due to censoring, but also reproduces other features of the actual curve such as asymmetry and a mean value closer to zero. This makes the censored quantile regression an appropriate estimator for bilateral migration stocks.

Figure 2 about here

We obtain the same kind of evidence if we compare distribution properties of the observed stock of migrants with those coming from PPML, OLS and Censored Quantile estimations - Table 3; where the 10^{th} , 50^{th} and 90^{th} percentiles of the observed distribution are very close to those estimated by our Censored Quantile estimator (and very

[27]The censored quantile regression density was estimated by using the Gaglianone and Lima's (2013) technique.

different from the ones estimated by OLS and PPML).

Table 3 about here

As mentioned in the introduction and section 3.2, we use the stocks rather than the flows of immigrants to cover a wider set of destination countries. Indeed, migration stocks from World Bank Bilateral Migration Matrix 2010 cover 200 destination countries while migration flows from OECD IMD cover only OECD destination countries. See Table 1 for the list of destination countries used in this paper. So we can analyse the specific effect of RTAs for developing destination countries. To this end, we include in our empirical specification an interaction term between the RTA dummy and a dummy being equal to one if the destination country is a developing country - as in the World Bank Classification.[28] The results reported in Table 4 show that RTAs stimulate with a bigger extent migration patterns towards developing countries. This is coherent with the idea that RTAs improve the information among potential migrants in origin countries. Indeed, the lack of information about the potential destinations is likely to be more important for developing destination countries. In the same spirit, we replicate this kind of specification but using a dummy for a south-south migration pattern (dummy equal to one if origin and destination countries are both developing). Similarly, we find that RTAs have stronger positive effect for south-south migration patterns than for other type of patterns. See columns (4)-(6) in Table 4.

4.2 Controlling for Endogeneity of RTAs

As argued above, the RTA dummy may suffer from a reverse causality problem leading to biased estimation of its coefficient. In order to fix this problem, we apply an IV moment condition approach to the censored quantile regression model. As in Kowalski (2013), our endogenous variable is discrete and the IV moment condition approach can be implemented by following the procedure developed by Chernozhukov and Hansen (2008). Kowalski

[28]Developing destination country dummy cannot be included since it is perfectly collinear with the fixed effects included in the regression

(2008) also showed that both control variable and moment condition approaches perform correctly with discrete endogenous variables.

Our instrument for the RTA dummy is the sum of RTAs that (respectively) origin and destination country has with the rest of the world. Indeed, Baldwin and Jaimovich (2012) show that the probability to have a RTA in common is positively related with the amount of RTAs that each country has with the rest of the world (to avoid trade diversion effects). This instrument is exogenous (valid) when RTAs with third countries do not affect bilateral migration settlement. This exclusion restriction can be tested by including our instrumental variable in the main specification. If the instrumental variable has significant effect on bilateral migration stock, the presence of RTAs with the rest of the World affects bilateral stocks and our instrument cannot be considered exogenous. We test the exclusion restriction in Table 7 finding null coefficient for our instrument. We may plausibly conclude that our instrument is exogenous and thus valid. The relevance of the instrumental variable is then based on first stage regression result, where our instrument is strongly correlated with the RTA dummy (coefficient on number of RTAs with Rest of the World - RoW - positive and significant). See Table 5.

Table 5 reports the estimation results that are robust against endogeneity of RTAs. When a RTA is signed, it affects bilateral migration flows and then the bilateral stock of migrants proportionally to the existing bilateral stock. So we do expect RTAs affecting more those country pairs having (already) high bilateral migration stocks: our censored quantile approach fits perfectly this need. Coherently with this intuition, in Table 5 we report estimation results for different quantiles of the migration distribution. Having a RTA in common matters for two quantiles of the distribution (50^{th} and 75^{th} quantiles) and does not matter for the 25^{th} quantile. Interestingly, Table 5 shows that the effect of RTA on the 50^{th} quantile (median) of the distribution is statistically lower than the effect on the 75^{th} quantile of the distribution, meaning that having a common RTA stimulates bilateral migration stocks by 33% at the median of the distribution and by 73% at the third quartile (75^{th} quantile). This suggests a strong asymmetric effect of RTAs along the distribution of the bilateral migration stocks.

Table 5 about here

To further explore such feature we use our robust estimator to show the RTA effect on various quantiles $\tau \in (0.1, 0.2, ..., 0.9)$ (results in Figure 3). The RTAs effect is monotonically increasing with the quantiles of the distribution: the higher the stock of bilateral migrants the stronger the effect of a reduction in migration costs through the signature of a RTA.

Figure 3 about here

Finally, we study the effect of the inclusion of visa and asylum provision in the RTAs using the IV moment condition approach (to addres the potential endogeneity of the visa-asylum dummy). The dummy standing for the presence of visa and asylum provision has been instrumented following the same approach as for the RTA dummy. Results reported in Table 6 show that RTAs still have their positive effect on migration stocks, with such effect magnified if the agreement includes visa and asylum provisions reducing the bureaucracy cost of migration. These last estimations stand out as a robustness check since they also include an EU dummy in the regression. Results in column (2) shows that having a common RTA stimulates bilateral migration stocks by 35% which increases up to 52% if the RTA includes also a provision on visa and asylum.

Table 6 about here

5 Concluding Remarks

This paper studies the effect of RTAs on the bilateral settlement of migrants; with the idea of capturing, through the signature of RTAs, the time varying bilateral migration costs.

To this end we estimated a structural gravity for migration model based on the seminal theoretical micro-foundations provided by Anderson (2011). The aim is to fill the gap between trade and migration empirical gravity estimation, by applying into a migration framework all the econometric advances made by recent literature on trade related gravity estimations (Santos Silva and Tenreyro 2006; Head and Mayer 2014; Baltagi, Egger and Pfaffermayr 2014).

First, we solve the zero migration flows problem by using the Powell's (1984, 1986) censored quantile regression, which, according to our results on migration, performs better than both OLS and PPML estimators. Second, we solve the endogeneity issue on RTA dummy by using the IV moment condition censored quantile regression approach. Using such new robust techniques we find two clear cut evidences: (i) RTAs stimulate bilateral migration settlements among member countries; (ii) the previous effect increases if the agreement includes provisions easing bureaucratic procedures on visa and asylum among member countries. We also find that RTA are particularly important for south-south migration pattern.

Although the main aim of this paper is to provide a practical toolkit for applied economists interested in estimating robust gravity models on migration, our paper suggests also interesting policy implications. RTAs might be used to regulate bilateral migration flows, and are informative for policy makers, who might use RTAs to increase migration inflows in the case of labour market shortages by leaving unchanged their migration policies. Finally, our results are particularly interesting for policy makers in developing countries. Indeed, RTAs represent for developing countries an opportunity to boost exports (and imports) but also - as showed by our results - a way to easy emigration toward (new) member countries.

6 References

Anderson, James, "The gravity model," *The Annual Review of Economics*, 3 (2011), 133-160.

Anderson, James and Eric van Wincoop, "Gravity with gravitas: a solution to the border puzzle," *American Economic Review*, 93 (2003), 170-192.

Baier, Scott L. and Jeffrey Bergstrand, "Do free trade agreements actually increase members' international trade?," *Journal of International Economics*, 71 (2007), 72-95.

Baldwin, Richard and Dany Jaimovich, "Are free trade agreements contagious?," *Journal of International Economics*, 88 (2012), 1-16.

Baltagi, H. Badi, Peter Egger and Michael Pfaffermayr, "Panel data gravity models of international trade," CESifo working paper n.4616 (2014).

Baker, Matthew, "Adaptive Markov chain Monte Carlo sampling and estimation in Mata," The Hunter College and the Graduate Center, working paper (2014).

Beine, Michel and Christopher Parsons "Climatic factors as determinants of International Migration," CESifo working paper n. 3747 (2012)
cm

Beine, Michel, Simone Bertoli and Jesus Fernández-Huertas Moraga, "A practitioners' guide to gravity models of international migration," forthcoming The World Economy (2015).

Bertoli, Simone and Jesus Fernández-Huertas Moraga, "Visa policies, networks and the cliff at the border," IZA Discussion Papers 7094 (2012).

Bertoli, Simone and Jesus Fernández-Huertas Moraga, "Multilateral resistance to migration," *Journal of Development Economics*, 102 (2013), 79-100.

Chernozhukov, Victor and Christian Hansen. "Instrumental variable quantile regression: a robust inference approach," *Journal of Econometrics*, (2008), 379-398.

Chernozhukov, Victor and Han Hong, "Three-Step censored quantile regression and extramarital affairs," *Journal of the American Statistical Association*, 97 (2002), 872-882.

Chernozhukov, Victor and Han Hong, "An MCMC approach to classical estimation," *Journal of Econometrics*, 115 (2003), 293-346.

Chernozhukov, Victor, Ivan Fernandez-Val, and Amanda Kowalski, "Quantile regression with censoring and endogeneity," NBER Working Papers, (2011), 169-197.

Figueiredo, Erik, Luiz Renato Lima and Georg Schaur, "Robust estimation of international trade specifications with heterogeneity in distance and policy effects," mimeo, (2014).

Gaglianone, Wagner and Luiz Renato Lima, "Constructing optimal density forecasts from point forecast combinations," forthcoming *Journal of Applied Econometrics*, (2013).

Galvao, Antonio, Carlos Lamarche, and Luiz Lima, "Estimation of censored quantile regression for panel data with fixed effects," *Journal of the American Statistical Association*, 108 (2013), 1075-1089.

Ghosh, Sucharita and Steven Yamarik, "Are regional trading arrangements trade creating?: An application of extreme bounds analysis," *Journal of International Economics*, 63 (2004), 369-395.

Grogger, Jeffrey and Gordon Hanson, "Income maximization and the selection and sorting of international migrants," *Journal of Development Economics*, 95 (2011), 42-57.

Hatton, Tim, "Explaining trends in UK immigration," *Journal of Population Economics*, 18 (2005), 719-740.

Head, Keith, Thierry Mayer and John Ries, "The erosion of colonial trade linkages after independence," *Journal of International Economics*, 81 (2010), 1-14.

Head, Keith and Thierry Mayer, "Gravity equations: workhorse, toolkit, and cookbook," chapter to appear in the Handbook of International Economics Vol. 4, eds. Gopinath, Helpman, and Rogoff, Elsevier. (2014).

Helpman, Elhanan, Marc Melitz and Yona Rubinstein, "Estimating trade flows: trading partners and trading volumes," *The Quarterly Journal of Economics*, 123 (2008), 441-487.

Horn, Henrik, Petros C. Mavroidis and André Sapir "Beyond the WTO? An Anatomy of EU and US Preferential Trade Agreements" *The World Economy*, 33 (2010), 1565-1588.

Koenker, Roger and Gilbert Bassett, Jr, "Regression quantiles," *Econometrica*, 46 (1978), 33-50.

Kowalski, Amanda, “Censored quantile instrumental variable estimates of the price elasticity of expenditure on medical care,” Massachsetts Institute of Technology dissertation, (2008), 1-96.

Kowalski, Amanda, “Censored quantile instrumental variable estimates of the price elasticity of expenditure on medical care,” mimeo, (2013).

Llull, Joan, “Understanding international migration: evidence from a new dataset of bilateral stocks (1960-2000),” Working Papers 715, Barcelona Graduate School of Economics, (2014).

Mayda, Anna, “International migration: a panel data analysis of the determinants of bilateral flows,” *Journal of Population Economics*, 23 (2010), 1249-1274.

Orefice, Gianluca, “International migration and trade agreements: the new role of PTAs,” *Canadian Journal of Economics*, (2015), 48(1).

Orefice, Gianluca and Nadia, Rocha "Deep integration and production networks: an empirical analysis?, *The World Economy*, (2014), 37(1), pp.106-136.

Ortega, Francesc and Giovanni Peri, “The Role of Income and Immigration Policies in Attracting International Migrants,” *Migration Studies*, 1 (2013), 47-74.

Özden, Caglar, Christopher Parsons, Maurice Schiff and Terrie Walmsley, “Where on earth is everybody? The evolution of global bilateral migration, 1960-2000,” *World Bank Economic Review*, 25 (2011), 12-56.

Powell, James, “Least absolute deviations estimation for the censored regression model,” *Journal of Econometrics*, 25 (1984), 303-25.

Powell, James “Censored regression quantiles,” *Journal of Econometrics*, 32 (1986), 143-55.

Ramos, Raul and Jordi Suriñach, “A gravity model of migration between ENC and EU,” IZA discussion paper number 7700, (2013).

Santos Silva, João and Silvana Tenreyro, “The log gravity,” *The Review of Economics and Statistics* 88 (2006), 641-658.

World Trade Report, "The WTO and the Preferential Trade Agreements: From Co-Existence to Coherence", Geneva: World Trade Organization (2011).

Table 1: List of Countries

Afghanistan	Dominican Republic	Lebanon	St. Lucia
Albania	Ecuador	Lesotho	St. Martin
Algeria	Egypt	Liberia	St. Pierre
Andorra	El Salvador	Libya	St. Vincent
Angola	Equatorial Guinea	Lithuania	Samoa
Antigua and Barbuda	Eritrea	Luxembourg	S. Tome Principe
Argentina	Estonia	Macedonia	Saudi Arabia
Armenia	Ethiopia	Madagascar	Senegal
Australia	Fiji	Malawi	Serbia
Austria	Finland	Malaysia	Seychelles
Azerbaijan	France	Maldives	Sierra Leone
Bahamas, The	French Guiana	Mali	Singapore
Bahrain	Gabon	Malta	Slovakia
Bangladesh	Gambia, The	Mauritania	Slovenia
Barbados	Georgia	Mauritius	Solomon Islands
Belarus	Germany	Mayotte	Somalia
Belgium	Ghana	Mexico	South Africa
Belize	Gibraltar	Micronesia	South Sudan
Benin	Greece	Moldova	Spain
Bermuda	Greenland	Mongolia	Sri Lanka
Bhutan	Grenada	Montenegro	Sudan
Bolivia	Guadeloupe	Montserrat	Suriname
Bosnia/Herzegovina	Guatemala	Morocco	Swaziland
Botswana	Guinea	Mozambique	Sweden
Brazil	Guinea-Bissau	Myanmar	Switzerland
Brunei	Guyana	Namibia	Syrian
Bulgaria	Haiti	Nauru	Taiwan
Burkina Faso	Honduras	Netherlands	Tajikistan
Burundi	Hong Kong	New Zealand	Tanzania
Cambodia	Hungary	Nicaragua	Thailand
Cameroon	Iceland	Niger	Timor-Leste
Canada	India	Nigeria	Togo
Cape Verde	Indonesia	Norway	Tonga
Cayman Islands	Iran	Oman	Trinidad and Tobago
Central African Rep.	Iraq	Pakistan	Tunisia
Chad	Ireland	Panama	Turkey
Chile	Israel	Papua N. Guinea	Turkmenistan
China	Italy	Paraguay	Uganda
Colombia	Jamaica	Peru	Ukraine
Comoros	Japan	Philippines	Unt. Arab Emirates
Congo, Rep.	Jordan	Poland	United Kingdom
Congo, Dem. Rep.	Kazakhstan	Portugal	United States
Costa Rica	Kenya	Puerto Rico	Uruguay
Cote d'Ivoire	Kiribati	Qatar	Uzbekistan
Croatia	Korea, Dem. Rep.	Reunion	Vanuatu
Cyprus	Korea, Rep.	Romania	Venezuela
Czech Republic	Kuwait	Russian Federation	Viet Nam
Denmark	Kyrgyzstan	Rwanda	Yemen
Djibouti	Lao PDR	St. Helena	Zambia
Dominica	Latvia	St. Kitts	Zimbabwe

Table 2: Baseline Estimations

	CQ Model[a] (1)	PPML (2)	OLS (3)	CQ Model[a] (4)	PPML (5)	OLS (6)	CQ Model[a] (7)	CQ Model[a] (8)
Agreement dummy								
RTA	0.057***	0.090	-0.032	0.077***	0.310***	-0.037	0.220***	0.243***
	(0.020)	(0.090)	(0.029)	(0.018)	(0.082)	(0.029)	(0.032)	(0.037)
Visa/Asylum	–	–	–	–	–	–	0.148***	0.079**
							(0.055)	(0.041)
Cost variables								
ln(GDP dest/ GDP orig)	0.115***	0.057	0.127***	0.314***	0.435***	0.318	0.224***	0.200***
	(0.010)	(0.074)	(0.018)	(0.008)	(0.012)	(0.334)	(0.011)	(0.023)
ln Pop. orig	0.699***	1.150***	0.526***	–	–	–	–	–
	(0.047)	(0.402)	(0.070)					
ln Pop. dest	-1.617***	-0.812***	-0.505***	–	–	–	–	–
	(0.023)	(0.023)	(0.078)					
WTO	0.226***	0.740***	0.145***	0.247***	0.578***	0.152***	0.077	0.144***
	(0.062)	(0.135)	(0.038)	(0.013)	(0.166)	(0.045)	(0.057)	(0.019)
ln Distance	-1.212***	-0.910***	-1.090***	-1.187***	-0.928***	-1.115***	-1.320***	-1.354***
	(0.060)	(0.053)	(0.015)	(0.006)	(0.047)	(0.015)	(0.007)	(0.094)
ln Stock 1960	0.564***	0..213***	0.375***	0.545***	0.213***	0.385***	0.340***	0.301***
	(0.015)	(0.013)	(0.005)	(0.015)	(0.012)	(0.004)	(0.009)	(0.039)
Border	1.765***	0.918***	1.329***	1.690***	0.884***	1.193***	0.885***	0.993***
	(0.038)	(0.098)	(0.067)	(0.037)	(0.090)	(0.068)	(0.054)	(0.088)
Colony	1.488***	1.180***	1.218***	1.522***	1.136***	1.179***	1.230***	0.914***
	(0.034)	(0.112)	(0.069)	(0.021)	(0.100)	(0.068)	(0.019)	(0.033)
Common Language	0.680***	0.167***	0.652***	0.574***	0.163***	0.643***	0.574***	0.500***
	(0.026)	(0.084)	(0.026)	(0.031)	(0.077)	(0.026)	(0.024)	(0.025)
EU	–	–	–	–	–	–	–	0.184***
								(0.044)
Country Effects	yes	yes	yes	no	no	no	no	no
Year Effects	yes	yes	yes	no	no	no	no	no
Country-by-Year FE	no	no	no	yes	yes	yes	yes	yes
Observations	80,345	80,345	43,907	80,345	80,345	43,907	16,126	16,126

Notes: (^{a}) model with $\tau = 0.50$. (***), (**) and (*) denote statistical significance at 1%, 5% and 10%, respectively. Standard errors in parentheses.

Table 3: Descriptive Statistics: Observed and Fitted Models[†]

	Observed	CQ Model^a	PPML	OLS
Min	0	0.0845	0.0142	0.0109
Max	1.16e+07	1.15e+07	1,06e+07	7,999,113
Mean	5796.99	60291.54	6055.91	3592.15
Total	4.44e+08	4.98e+08	4.44e+08	2.25e+08
Centile 0.10	0	0.6008	9.6663	1.976
Centile 0.50	5	7.63	247.60	30.19
Centile 0.90	2,568	1,707.59	7,137.99	1,198.76
Recovered Values^b	0	6,763,604	3.60e+07	0

Notes: (†) models from Table 2: columns (3)-(5). (^{a}) model with $\tau = 0.50$. (^{b}) Estimates for the migration stock when observed values are equal to zero.

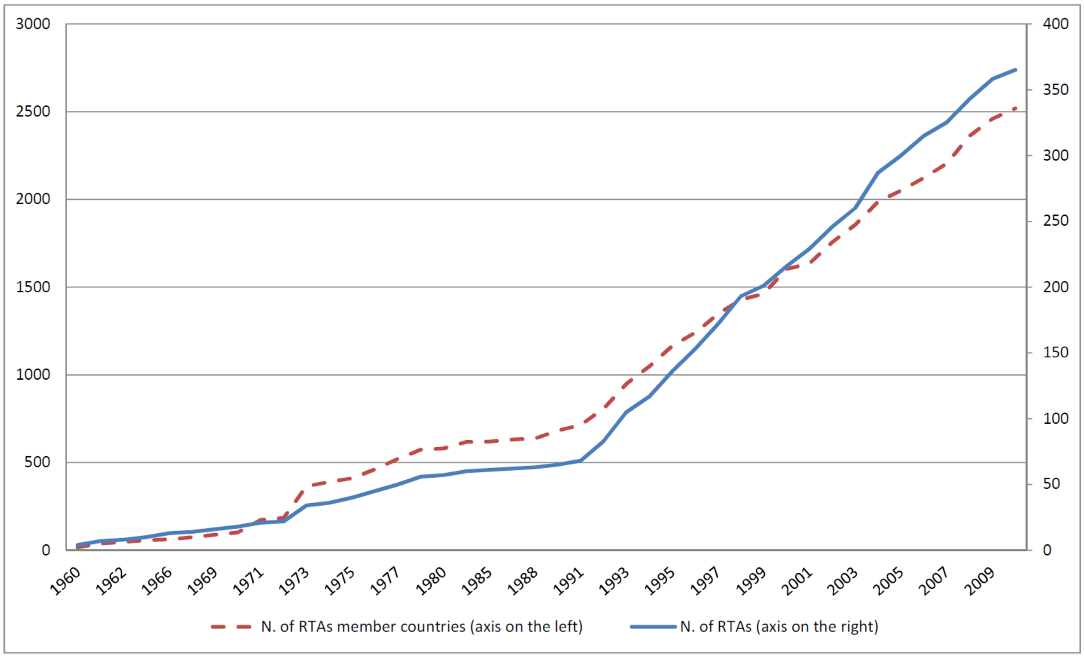


Figure 1: Number of RTAs and member countries over the period 1996-2010
Note: In the count of RTAs' member countries, if a country has n RTA, it is counted n times.

Table 4: Extended specifications: developing and south-south interactions.

	CQ Model[a] (1)	PPML (2)	OLS (3)	CQ Model[a] (4)	PPML (5)	OLS (6)
Agreement dummy						
RTA	0.084*	0.090	-0.023	0.086***	0.203**	-0.093**
	(0.043)	(0.082)	(0.038)	(0.021)	(0.089)	(0.036)
RTA×Developing destinations	0.057***	0.763***	-0.031	–	–	–
	(0.016)	(0.186)	(0.034)			
RTA×South-south	–	–	–	0.296***	0.425**	0.243***
				(0.028)	(0.192)	(0.057)
South-south	–	–	–	-0.334***	-0.567***	-0.381***
				(0.058)	(0.133)	(0.033)
Cost variables						
ln(GDPDest/ GDPOrig)	0.317***	0.547***	0.320	0.351***	0.601***	0.313
	(0.012)	(0.016)	(0.334)	(0.012)	(0.124)	(0.328)
WTO	0.277***	0.516***	0.152***	0.283***	0.571***	0.174***
	(0.024)	(0.168)	(0.046)	(0.015)	(0.162)	(0.046)
ln Distance	-1.199***	-0.948***	-1.114***	-1.182***	-0.951***	-1.133***
	(0.007)	(0.040)	(0.015)	(0.004)	(0.050)	(0.015)
ln Stock 1960	0.549***	0.215***	0.385***	0.540***	0.217***	0.386***
	(0.016)	(0.011)	(0.004)	(0.011)	(0.012)	(0.004)
Border	1.392***	0.854***	1.195***	1.687***	0.890***	1.195***
	(0.044)	(0.088)	(0.068)	(0.014)	(0.090)	(0.067)
Colony	1.344***	1.131***	1.180***	1.510***	1.102***	1.163***
	(0.031)	(0.097)	(0.068)	(0.034)	(0.100)	(0.068)
Common Language	0.611***	0.154**	0.643***	0.701***	0.149**	0.642***
	(0.022)	(0.077)	(0.026)	(0.024)	(0.078)	(0.026)
Country-by-Year FE	yes	yes	yes	yes	yes	yes
Observations	80,345	80,345	43,907	80,345	80,345	43,907

Notes: ([a]) model with $\tau = 0.50$. (***), (**) and (*) denote statistical significance at 1%, 5% and 10%, respectively. Standard errors in parentheses.

Table 5: Gravity Models with Instrumental Variables

	First Stage: OLS		
Dependent variable: RTA			
N. of RTAs with RoW	0.039***		
	(0.001)		
WTO	0.102***		
	(0.004)		
ln(GDPDest/GDPOrig)	0.234		
	(0.056)		
ln Distance	-0.238***		
	(0.002)		
ln Stock 1960	0.209***		
	(0.006)		
Border	0.894***		
	(0.016)		
Colony	1.126***		
	(0.237)		
Common Language	0.761		
	(0.099)		
Second Stage: Censored Quantile Models			
Specifications	$\tau = 0.25$	$\tau = 0.50$	$\tau = 0.75$
RTA[a]	-0.001	0.289***	0.548***
	(0.021)	(0.013)	(0.035)
ln(GDPDest/GDPOrig)	0.245***	0.394***	0.409***
	(0.123)	(0.102)	(0.124)
WTO	0.274***	0.218***	0.203***
	(0.043)	(0.055)	(0.043)
ln Distance	-0.863***	-0.743***	-0.709***
	(0.034)	(0.044)	(0.045)
ln Stock 1960	0.543***	0.554***	0.574***
	(0.120)	(0.042)	(0.098)
Border	1.165***	0.903***	0.809***
	(0.232)	(0.345)	(0.235)
Colony	1.546***	1.654***	1.667***
	(0.327)	(0.334)	(0.369)
Common Language	0.321***	0.398***	0.421***
	(0.092)	(0.088)	(0.109)
Country-by-Year Effects	yes	yes	yes
Observations	80,345	80,345	80,345

Notes: ([a]) Controlled for endogeneity. (***), (**) and (*) denote statistical significance at 1%, 5% and 10%, respectively. Standard errors in parentheses.

Table 6: Gravity Models with Instrumental Variables: Visa and Asylum

	Censored Quantile Models	
Specifications[a]	(1)	(2)
Agreement dummies		
RTA[b]	0.299***	0.300***
	(0.054)	(0.084)
Visa/Asylum[b]	0.213***	0.123***
	(0.043)	(0.033)
EU	–	0.055***
		(0.013)
ln(GDPDest/GDPOrig)	0.214***	0.244***
	(0.056)	(0.023)
WTO	0.054	0.124***
	(0.043)	(0.050)
ln Distance	-1.004***	-0.955***
	(0.004)	(0.064)
ln Stock 1960	0.301***	0.287***
	(0.078)	(0.088)
Border	0.817***	0.821***
	(0.092)	(0.123)
Colony	1.012***	1.006***
	(0.234)	(0.223)
Common Language	0.540***	0.521***
	(0.084)	(0.096)
Country-by-Year Effects	yes	yes
Observations	16,126	16,126

Notes: ([a]) Those specifications consider $\tau = 0.50$. ([b]) Controled by endogeneity. (***), (**) and (*) denote statistical significance at 1%, 5% and 10%, respectively. Standard errors in parentheses.

Table 7: Diversion Effect

	Censored Quantile Models		
Specifications	$\tau = 0.25$	$\tau = 0.50$	$\tau = 0.75$
	(1)	(2)	(3)
Agreement dummies			
RTA	0.111**	0.166***	0.193***
	(0.055)	(0.032)	(0.032)
N. of RTAs with RoW	-0.018	0.002	0.033
	(0.054)	(0.014)	(0.043)
Cost variables			
ln(GDPDest/GDPOrig)	0.388***	0.395***	0.404***
	(0.132)	(0.139)	(0.177)
WTO	0.383***	0.389***	0.423***
	(0.049)	(0.044)	(0.057)
ln Distance	-1.510***	-1.425***	-1.380***
	(0.205)	(0.192)	(0.198)
ln Stock 1960	0.243*	0.266***	0.268***
	(0.010)	(0.089)	(0.084)
Border	0.958***	1.000***	1.120***
	(0.397)	(0.383)	(0.367)
Colony	1.011***	1.169***	1.169***
	(0.233)	(0.312)	(0.362)
Common Language	0.896***	1.010***	1.176***
	(0.255)	(0.268)	(0.322)
Country-by-Year Effects	yes	yes	yes
Observations	80,345	80,345	80,345

Notes: (***), (**) and (*) denote statistical significance at 1%, 5% and 10%, respectively. Standard errors in parentheses.

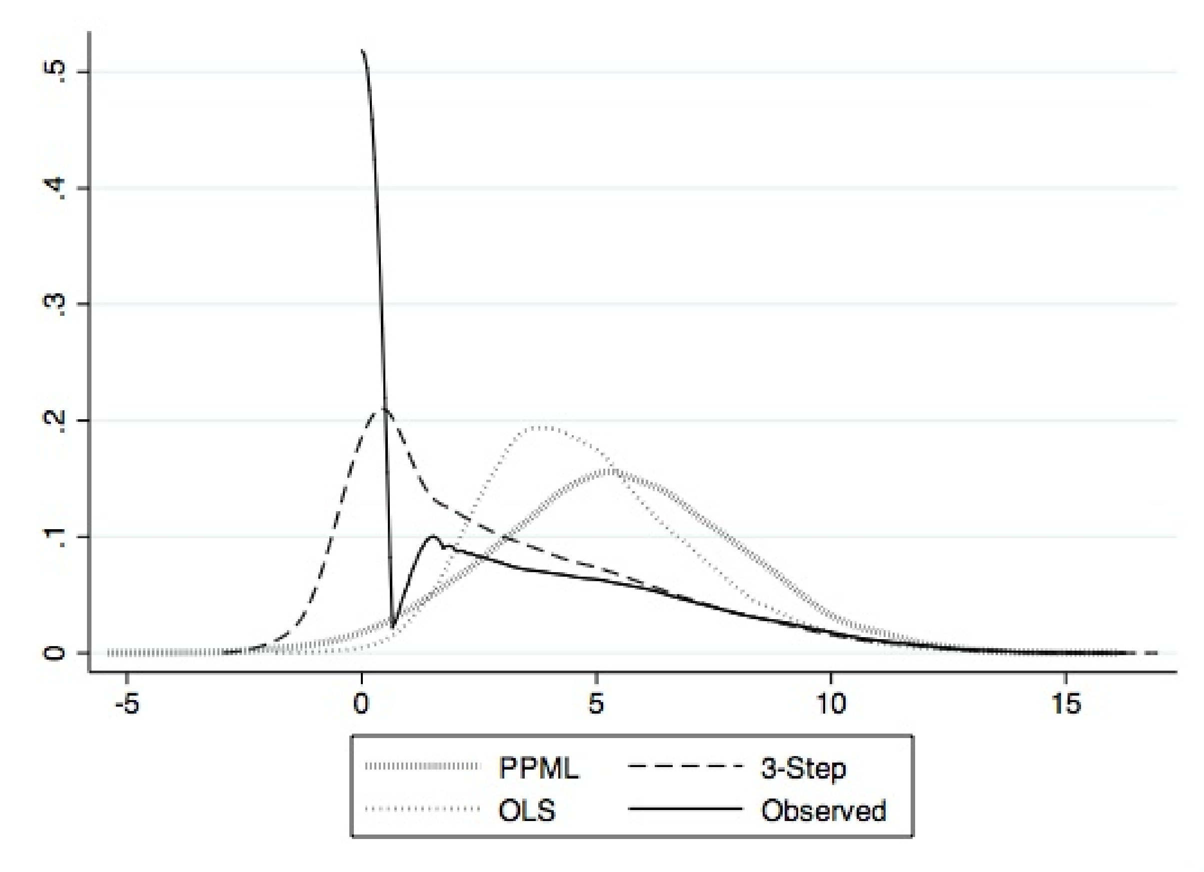


Figure 2: Observed and Fitted Models. Log of migration stock is reported in the horizontal axis. Density in the vertical axis.

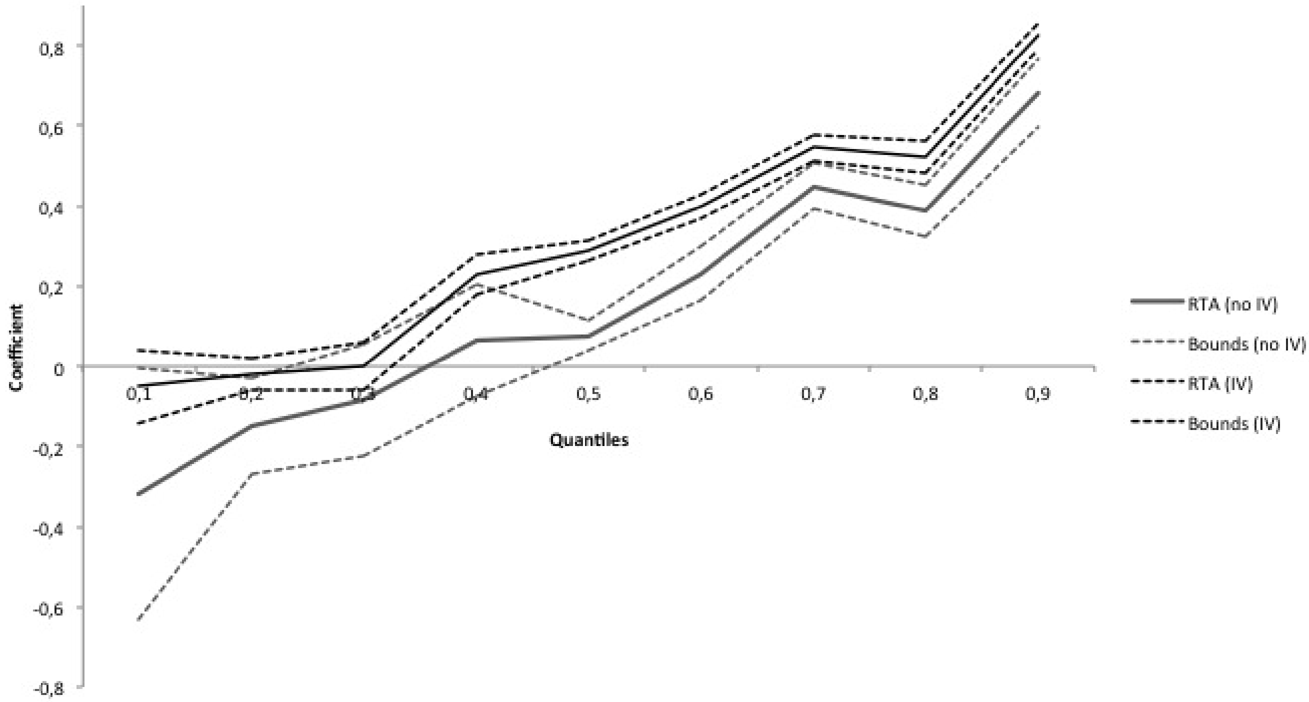


Figure 3: RTA Effects Across Quantiles – Country-by-Year fixed effects estimation.

A Appendix

A.1 Using Bilateral Migration Flows

Table A.1: Dependent variable: bilateral migration flows

	CQ Model[a] (1)	PPML (2)	OLS (3)	CQ Model[a] (4)	PPML (5)	OLS (6)
Agreement dummy						
RTA	0.255***	0.339***	0.244***	0.264***	0.656***	0.434***
	(0.052)	(0.068)	(0.039)	(0.024)	(0.091)	(0.086)
Cost variables						
ln(GDPDest/GDDOrig)	0.205***	0.234***	0.202***	0.194***	0.334***	0.197***
	(0.056)	(0.073)	(0.026)	(0.036)	(0.060)	(0.028)
ln Pop. orig	1.908***	9.012***	3.671***	–	–	–
	(0.032)	(1.105)	(0.746)			
ln Pop. dest	-0.782***	-5.578***	-2.099***	–	–	–
	(0.045)	(0.704)	(0.410)			
ln Distance	-0.369***	-0.498***	-0.725***	-0.544***	-0.541***	-0.691***
	(0.056)	(0.070)	(0.040)	(0.083)	(0.051)	(0.041)
ln Stock 1991	0.342***	0.511***	0.233***	0.273***	0.492***	0.232***
	(0.065)	(0.020)	(0.006)	(0.063)	(0.017)	(0.042)
Border	0.047	-0.177***	-0.177***	-0.104***	-0.123**	-0.126**
	(0.097)	(0.068)	(0.058)	(0.090)	(0.065)	(0.060)
Colony	0.220***	0.837***	0.806***	0.165***	0.919***	0.822***
	(0.041)	(0.075)	(0.060)	(0.036)	(0.059)	(0.061)
Common Language	0.111***	0.466***	0.646***	0.105***	0.553***	0.615***
	(0.087)	(0.071)	(0.056)	(0.021)	(0.067)	(0.057)
Country Effects	yes	yes	yes	no	no	no
Year Effects	yes	yes	yes	no	no	no
Country-by-Year FE	no	no	no	yes	yes	yes
Observations	7,589	7,589	5,807	7,589	7,589	5,807

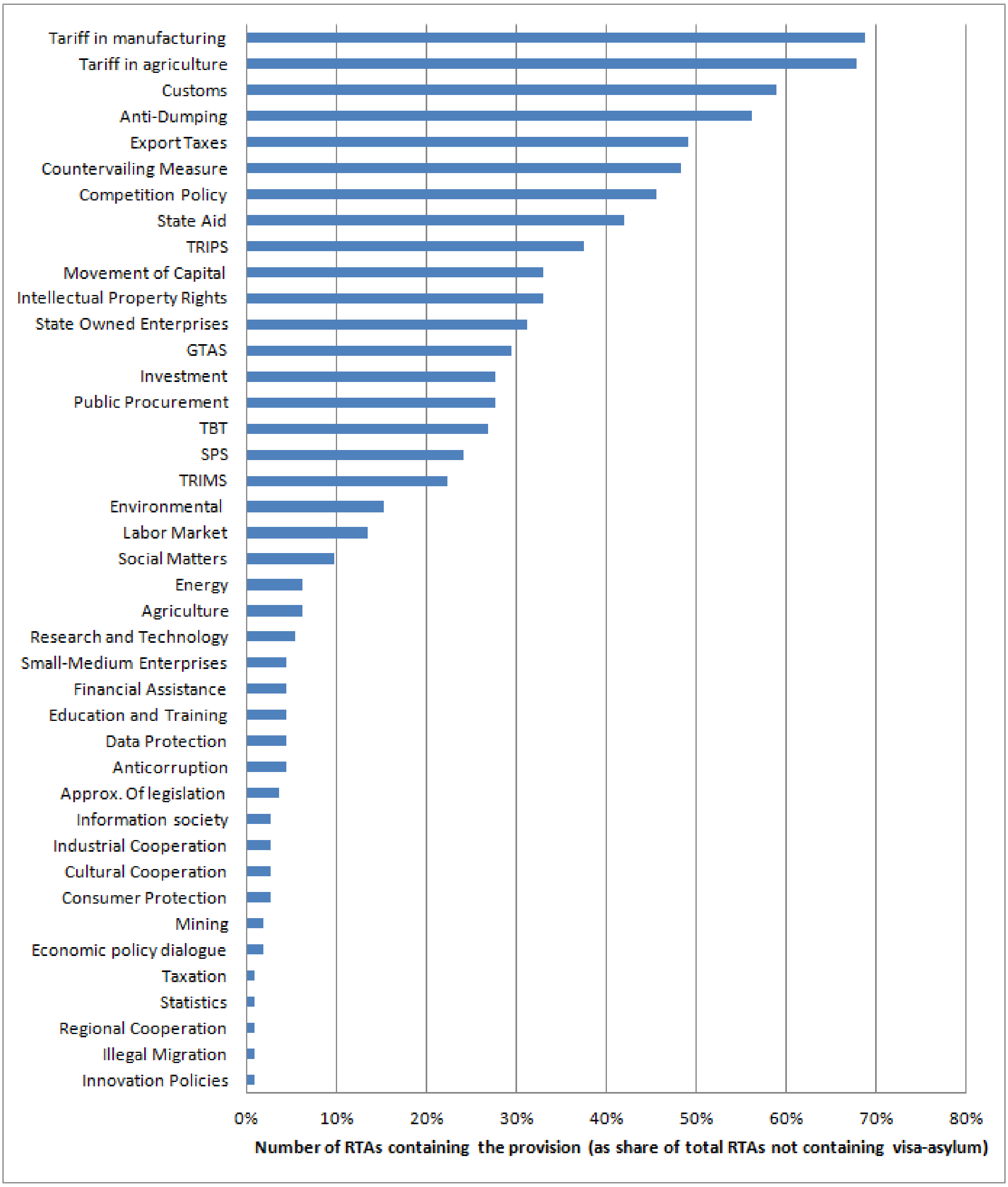
Notes: (^{a}) model with $\tau = 0.50$. (***), (**) and (*) denote statistical significance at 1%, 5% and 10%, respectively. Standard errors in parentheses.

Table A.2: Extended specifications: PTA/FTA and CU.

	CQ Model[a]
	(1)
Agreement dummy	
PTA/FTA	0.064***
	(0.023)
CU	0.336***
	(0.015)
Cost variables	
ln(GDPDest/ GDPOrig)	0.321***
	(0.012)
WTO	0.264***
	(0.027)
ln Distance	-1.184***
	(0.008)
ln Stock 1960	0.549***
	(0.016)
Border	1.358***
	(0.058)
Colony	1.358***
	(0.038)
Common Language	0.602***
	(0.028)
Country-by-Year FE	yes
Observations	80,345

Notes: ([a]) model with $\tau = 0.50$. (***), (**) and (*) denote statistical significance at 1%, 5% and 10%, respectively. Standard errors in parentheses.

Figure A1: Provisions included in RTAs that do not contain visa-and-asylum provision.



Source: Authors' calculations on the WTO database

A.2 Econometric Model

In this section, we provide technical details on the econometric modelling used in the paper.

Figueiredo, Lima and Schaur (2014) considered the following generalization of model (5)

$$
\begin{aligned}
M_{ij} &= \exp\left(x_{ij}\beta\right)\eta_{ij}, \qquad \text{(A-1)}\\
\eta_{ij} &= \exp\left[\left(x_{ij}\gamma\right)\varepsilon_{ij}\right],\\
\varepsilon_{ij} &\sim i.i.d.F_{\varepsilon}\left(\mu,\sigma^{2}\right).
\end{aligned}
$$

$F_{\varepsilon}\left(\cdot\right)$ is an unknown continuous distribution function of ε_{ij}, where $F_{\varepsilon}^{-1}\left(\tau\right) = Q_{\tau}\left(\varepsilon_{ij}\right)$ is the $\tau - th$ quantile of ε_{ij} and $\tau \in (0,1)$. Let $Q_{\tau}\left(M_{ij}|x_{ij}\right)$ denote the the $\tau - th$ conditional quantile of M_{ij}. Thus, the quantiles of M_{ij} can be written as

$$
Q_{\tau}\left(M_{ij}|x_{ij}\right) = \exp\left(x_{ij}\beta\left(\tau\right)\right), \qquad \text{(A-2)}
$$

where $\beta + \gamma \cdot Q_{\tau}\left(\varepsilon_{ij}\right)$. For instance, when $\tau = 0.5$, $Q_{\tau}\left(\varepsilon_{ij}\right)$ becomes $Median\left(\varepsilon_{ij}\right)$ and $\beta\left(0.5\right) = \beta + \gamma \cdot Median\left(\varepsilon_{ij}\right) = \beta_{median}$.

Now, by the property of equivariance of the quantile function, if one assumes that $Median\left(\eta_{ij}|x_{ij}\right) = 1$, then this will imply that $Median\left(\varepsilon_{ij}\right) = 0$, then $\beta_{median} = \beta$ and $Median\left(M_{ij}|x_{ij}\right) = \exp\left(x_{ij}\beta\right)$. Therefore, the median estimator identifies the parameter β.[29]

The important consequence of $Median\left(\eta_{ij}|x_{ij}\right) = 1$ implying $Median\left(\varepsilon_{ij}\right) = 0$ is that, unlike the OLS or PPML based estimation, the identification of quantiles in the exponential model leads to the identification of quantiles in the log-linear model and vice-versa.[30] In other words, for any $\tau \in (0,1)$, equivariance gives $Q_{\tau}\left(\ln(M_{ij})|x_{ij}\right) = \ln\left[Q_{\tau}\left(M_{ij}|x_{ij}\right)\right] = \ln\left[\exp\left(x_{ij}\beta\left(\tau\right)\right)\right] = x_{ij}\beta\left(\tau\right)$, where $\beta\left(\tau\right) = \beta + \gamma \cdot Q_{\tau}\left(\varepsilon_{ij}\right)$. Because the quantile approach

[29] The median estimator is just a special case of the quantile regression estimator proposed by Koenker and Bassett (1978).

[30] Figueiredo, Lima and Schaur (2014) conducted Monte Carlo simulations to assess the bias of the PPML when its identifying condition does not hold.

identifies both, the exponential and the log linearized model, we can focus our estimation on the log-linear model to avoid the problem of overweighing large bilateral migration stocks.[31]

However, as in the trade literature, the main problem related to the log-linear model (6) is a large amount of zero migration flows (see Orefice 2015, Ramos and Suriñach 2013, and among others).[32] For these observations, taking the log of zero will automatically lead the computer to drop them. To address this issue, we follow Cameron and Trivedi (2009) and estimate a Tobit for lognormal model.

In our notation, this solution corresponds to set the observed value to $\ln(z_{ij}) = \max(0, \ln(M_{ij}))$ and, therefore, $\ln(z_{ij})$ is set equal to 0 whenever the original observations are subject to statistical rounding, that is, whenever $M_{ij} < 1$, where 1 is the minimum uncensored value of migration stocks.

However, a Tobit model is used to estimate the conditional mean function and relies strongly on the assumptions of normality and homoskedasticity. The approach proposed in this paper can be seen a generalization of the Tobit model in the sense that it allows us to identify the quantiles of the conditional distribution of $\ln(M_{ij})$ (and M_{ij}) and does not rely on any distributional assumption such as normality and homoskedasticity. Given that $\ln(z_{ij}) = \max(0, \ln(M_{ij}))$, the equivariance property naturally leads to the censored quantile regression model

$$\begin{aligned} Q_\tau[\ln(z_{ij})|x_{ij}] &= \max(0, Q_\tau[\ln(M_{ij})|x_{ij}]) \\ &= \max(0, x_{ij}\beta(\tau)). \end{aligned} \tag{A-3}$$

The censored quantile model (A-3), devoloped by Powell (1984, 1986), provides a way to do valid inference in Tobin-Amemiya models without distributional assumptions and with

[31] Notice that we no longer have the bias problem associated to the log-linear model that was previously pointed out by Santos Silva and Tenreyro (2006).

[32] This problem has been discussed in the trade literature since the early 1980s. See, for instance, Head and Mayer (2014, section 5.2).

heteroskedasticity of unknown form. The Powell's censored quantile regression is defined to maximize the objective function:

$$L_n(\beta) = -\sum_{i,j=1}^{n} w_{ij}\rho_\tau[\ln(z_{ij}) - \max(0, x_{ij}\beta(\tau))], \tag{A-4}$$

where ρ_τ represents the traditional loss function of quantile regression developed by Koenker and Bassett (1978), w_{ij} is a weight. Chernozhukov and Hong (2003) show that the extremum estimator represented by (A-4) has optimization problems caused by the nonconvexity of the objective function. A robust solution to optimize this function is provided by Chernozukov and Hong (2003), in which the authors use the Markov chain Monte Carlo (MCMC) method to estimate a pseudo-quadratic objective function such as (A-4).

Nevertheless, as shown in the subsection 2.1, a successful estimator for the log linear gravity model for migration must address the endogeneity issue. This is carried out by applying the IV moment condition approach developed by Chernozhukov and Hansen (2008), which consist in adding to equation (A-4) an additional pre-step to handle endogeneity.[33]

In other words, let $x_{ij} = (x_{0,ij}, x_{1,ij})$ where $x_{0,ij}$ is the endogenous variable while $x_{1,ij}$ are exogenous, $b(\tau) = (\alpha_0, \beta_0)$ are the corresponding parameters for a given τ, and $v_{ij} = (x_{1,ij}, v_{0,ij})$, where $v_{0,ij}$ is a set of instrumental variables. Then, the estimating procedure goes as follows:

1. consider $\ln(z_{ij}) - x_{0,ij} \cdot \alpha_0 = x_{1,ij} \cdot \beta_0 + v_{0,ij} \cdot \zeta + \epsilon_{ij}$, with $Q_\tau[\epsilon_{ij}|v_{0,ij}, x_{1,ij}] = 0$. For a given value of $\alpha_k \in (\alpha_1, ..., \alpha_J)$, we run a quantile regression of $\ln(z_{ij}) - x_{0,ij}^T \cdot \alpha_k$ on $(x_{1,ij}, v_{0,ij})$, and compute the Wald statistic (W_n) corresponding to the test of $\zeta(\alpha_j) = 0$. Then we define the estimator of α_0 as $\hat{\alpha} = \arg\min_{k=1,..,J} W_n(\alpha_k)$. In this paper we assumed that, for a given quantile τ, there is a searching grid with 200 values of α_k.

[33] Alternatively, we should consider the three-step estimator provided by Chernozhukov and Hong (2002). However, this approach uses the control function approach to endogeneity, and the assumptions necessary for the control function approach are less likely to be satisfied when the endogenous variable is discrete (see Kowalski 2013).

2. define $\tilde{x}_{ij} = (x_{0,ij} \cdot \hat{\alpha}, x_{1,ij})$ and replace it into the equation (A-4):

$$L_n(\beta) = -\sum_{i,j=1}^{n} w_{ij}\rho_\tau[\ln(z_{ij}) - \max(0, \tilde{x}_{ij}\beta(\tau))], \tag{A-5}$$

Finally, the vector of parameters from (A-5) is estimated by using the MCMC algorithm developed by Chernozukov and Hong (2003).