

Trade facilitation and country size

Mohammad Amin · Jamal Ibrahim Haidar

Received: 2 October 2012 / Accepted: 25 October 2013 / Published online: 11 December 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract It is argued that compared with large countries, small countries rely more on trade and therefore are more likely to adopt liberal trading policies. The present paper extends this idea beyond the conventional trade openness measures by analyzing the relationship between country size and the number of documents required to export and import, a measure of trade facilitation. Three important results follow. First, trade facilitation does improve as country size becomes smaller; that is, small countries perform better than large countries in terms of trade facilitation. Second, the relationship between country size and trade facilitation is nonlinear, much stronger for the relatively small than the large countries. Third, contrary to what existing studies might suggest, the relationship between country size and trade facilitation does not appear to be driven by the fact that small countries trade more as a proportion of their gross domestic product than the large countries.

Keywords Country size · Trade facilitation · Openness

JEL Classification F10 · F13 · F14 · F15 · F50 · F55 · H10 · K00 · L5 · O20 · O24

1 Introduction

There is a large body of work that shows that small countries are likely to benefit more from international trade or a liberal trading regime than large countries. Small-

M. Amin
Enterprise Analysis Unit, World Bank, Washington, DC 20433, USA
e-mail: mamin@worldbank.org

J. I. Haidar (✉)
Paris School of Economics, University of Paris-1, Pantheon Sorbonne,
Paris 75014, France
e-mail: Jamal.Haidar@parisschoolofeconomics.eu

ness of market limits the exploitation of economies of scale, forcing the relatively small countries to expand market size through international trade beyond their political borders (Alesina 2002; Alesina and Wacziarg 1998). However, most of the evidence on the relationship between trade openness and country size is largely focused on measures of trade openness that include trade volume (exports plus imports as percentage of GDP) and border taxes (tariffs). There is almost no evidence on how country size affects trade facilitation at the microlevel, the focus of the present paper. For example, for a sample of over 80 countries and controlling for a number of other determinants of trade openness, Alesina and Wacziarg (1998) find that doubling country size as measured by total population is associated with a 9 % point reduction in the trade-to-GDP ratio. Qualitatively similar results are also reported for macro-level trade policy measures including tariff rates. Similar findings for the exports plus imports-to-GDP ratio are also reported in Rose (2006) and Alesina and Spolaore (2003).

The present paper extends the literature discussed above in two important ways. First, in contrast to existing studies, the present paper focuses on micro-level trade facilitation or “inside the border” measures to define how liberal the trading regime is. With the decline in tariff and non-tariff barriers, greater attention is now being devoted to trade facilitation measures. However, the relationship between trade facilitation and country size is still unexplored. Second, the paper highlights a strong nonlinearity in how country size affects trade facilitation. Consistent with the broader literature on trade and country size mentioned above, we find that trade facilitation becomes worse as country size increases. However, this relationship is much stronger for the relatively small countries and weaker for the large ones. The implications of this nonlinear relationship for the broader literature are discussed.

The motivation for focusing on trade facilitation measures comes from existing studies that show that with the decline of tariff and non-tariff barriers around the globe, it is the trade facilitation measures that are becoming increasingly more important for the overall expansion of international trade (see, for example, Wilson et al. 2003). There is no standard definition of trade facilitation in public policy discourse. In a narrow sense, trade facilitation efforts simply address the logistics of moving goods through ports and the documentation associated with cross-border trade. Some of the factors included under trade facilitation are Internet availability (Freund and Weinhold 2002), time to clear shipments at ports (Djankov 2010), and standards harmonization and automated customs procedures (Hertel et al. 2001). The present paper focuses on a regulatory aspect of trade facilitation, the number of documents required to export and import as measured by the World Bank Doing Business project. There is a lot of variation in the measure across countries. For example, as of May 2010, while two documents are required to export a container from France, it requires 11 documents to do the same in Namibia. Is some of this cross-country variation in the number of required documents due to differences in country size? The present paper attempts to answer this question.

Much like the studies mentioned above, our focus is on a single element of trade facilitation. We would like to caution that trade facilitation is a multi-faceted phenomenon and there is no guarantee that the results presented in this paper for the number

of documents required to export and import would necessarily extend to other aspects of trade facilitation. A careful analysis of the various trade facilitation measures one by one may be required to understand the determinants and effects of these different measures.

The motivation for exploring the possibility of nonlinearity in the relationship between country size and trade facilitation is largely empirical in nature. That is, there is no strong theoretical reason to expect the stated relationship to be nonlinear or linear. Without a formal theoretical underpinning, it is difficult for us to establish with any reasonable degree of confidence exactly how and why the nonlinearity exists, what are the underlying causes of this nonlinearity, and whether we should expect similar nonlinearity with respect to other determinants of trade facilitation measures or not. At present, we can only speculate on the possible explanations of the stated nonlinearity. For example, one could argue that, at the margin, economies of scale may be most pressing when country size is small to begin with. Hence, the country size and trade facilitation relationship is likely to be stronger among the relatively small than the large countries. However, this is merely speculative and needs to be empirically validated or rejected. We hope that findings in the present paper will motivate other researchers to explore the nonlinearity in more detail, something that is beyond the scope of the present paper.

Above, we mentioned some studies related to international trade that use population as a proxy for country and market size. However, the practice of using population as a proxy for country/market size is not restricted to international trade. For example, see [Schwarz \(2007\)](#) on how country size measured by the population of the country affects the level corporate tax burden; [Root \(1999\)](#) on how small country size as measured by population limits economies of scale, forcing the government to be more efficient and therefore less corrupt; and [Rose \(2006\)](#) on how population as a proxy for country size correlates with a number of macroeconomic variables including among others, life expectancy at birth, literacy rate, political rights, and inflation.

The plan of the remaining sections is as follows. In Sect. 2, we describe the data and the empirical methodology. Regression results for our main specification along with a number of robustness checks are discussed in Sect. 3. To raise our confidence against possible endogeneity concerns with our main results, instrumental variable regression results are provided in Sect. 4. The concluding section summarizes our main findings and suggests scope for future work.

2 Methodology and data description

We focus on the developing countries which include all the low and middle income countries as defined by the World Bank and for which data are available for our main variables of interest. There are 106 such countries in our sample (listed in Table 1). Our main regression results are based on the ordinary least squares (OLS) methodology. Instrumental variables (IV) regression results are also reported in detail in order to increase our confidence against a possible endogeneity problem with our main estimation results. All regression results use Huber–White robust standard errors and have

Table 1 List of countries included in the sample

Albania	Eritrea	Nicaragua
Algeria	Ethiopia	Niger
Angola	Gabon	Nigeria
Antigua and Barbuda	Georgia	Pakistan
Argentina	Ghana	Papua New Guinea
Armenia	Grenada	Paraguay
Azerbaijan	Guatemala	Peru
Bangladesh	Guinea	Philippines
Belarus	Guinea-Bissau	Romania
Belize	Guyana	Russian Federation
Benin	Honduras	Rwanda
Bhutan	India	Senegal
Bolivia	Indonesia	Seychelles
Bosnia and Herzegovina	Iran, Islamic Republic	South Africa
Botswana	Jamaica	Sri Lanka
Brazil	Jordan	St. Kitts and Nevis
Bulgaria	Kazakhstan	St. Lucia
Burkina Faso	Kenya	St. Vincent and the Gren.
Burundi	Kyrgyz Republic	Sudan
Cambodia	Lao PDR	Swaziland
Cameroon	Lebanon	Syrian Arab Republic
Cape Verde	Lesotho	Tajikistan
Central African Republic	Lithuania	Tanzania
Chad	Macedonia, FYR	Thailand
Chile	Madagascar	Togo
China	Malawi	Tunisia
Colombia	Malaysia	Turkey
Congo, Republic	Mali	Uganda
Costa Rica	Mauritania	Ukraine
Côte d'Ivoire	Mauritius	Uruguay
Djibouti	Moldova	Vanuatu
Dominica	Mongolia	Venezuela, R.B.
Dominican Republic	Morocco	Vietnam
Ecuador	Mozambique	Zambia
Egypt, Arab Republic	Namibia	
El Salvador	Nepal	

been checked for possible outliers that may have unduly large effects on the main regression results.

Table 2 formally defines all the variables used in the paper along with data sources. Table 3 provides descriptive statistics of the variables, and Table 4 shows the correlation between them.

Table 2 Description of variables

Documents	All documents required per shipment to export and import the goods are recorded. It is assumed that the contract has already been agreed upon and signed by both parties. Documents required for clearance by government ministries, customs authorities, port and container terminal authorities, health and technical control agencies, and banks are taken into account. Since payment is by letter of credit, all documents required by banks for the issuance or securing of a letter of credit are also taken into account. Documents that are renewed annually and that do not require renewal per shipment (for example, an annual tax clearance certificate) are not included. We take the average value of the variable over all years for which data are available and use log values of the average. Source: Doing Business, World Bank
Population	Total population of a country, averaged over 2001–2005. Log values are used. Source: World Development Indicators, World Bank
Population ²	Square of Population
Income	GDP per capita (PPP adjusted and at constant 2005 USD). We take the average value of the variable over 2001–2005 and then the log of the average values. Source: World Development Indicators, World Bank
Trade-to-GDP ratio	Volume of exports plus imports as a percentage of GDP. We take the average over the 2001–2005 period and then the log of the average values. Source: World Development Indicators, World Bank
Weighted tariff	Tariff rate weighted by the bilateral volume of trade for the concerned product. The variable is defined for all products and we use the log of the average value where the average is taken over the 2001–2005 period. Source: World Development Indicators, World Bank
Latitude	Absolute distance from the equator. Source: La Porta et al. (1999)
Regulation (EODB)	Ease of Doing Business Index. We use average values taken over all years for which data are available. Source: Doing Business, World Bank
Common law	Dummy for the English Common law. Source: La Porta et al. (1999)
Socialist law	Dummy for the Socialist law. Source: La Porta et al. (1999)
Catholic	Dummy indicating if the majority of population is Catholic. Source: La Porta et al. (1999)
Muslim	Dummy indicating if the majority of population is Muslim. Source: La Porta et al. (1999)
Protestant	Dummy indicating if the majority of population is Protestant. Source: La Porta et al. (1999)
Ethnic	A measure of ethnic fractionalization. Higher values imply more ethnic fractionalization or diversity. Source: Alesina and Spolaore (2003) , <i>Journal of Economic Growth</i> , June 2003; Table A1
Area	Log of total land area of the country in square kilometer as of 2005. Source: World Development Indicators, World Bank
Time	The time for exporting and importing is recorded in calendar days. The time calculation for a procedure starts from the moment it is initiated and runs until it is completed. If a procedure can be accelerated for an additional cost and is available to all trading companies, the fastest legal procedure is chosen. Fast-track procedures applied to firms located in an export processing zone are not taken into account because they are not available to all trading companies. Ocean transport time is not included. It is assumed that neither the exporter nor the importer wastes time and that each commits to completing each remaining procedure without delay. Procedures that can be completed in parallel are measured as simultaneous. The waiting time between procedures—for example, during unloading of the cargo—is included in the measure. We take average value of the variable over all years for which data are available and use log values of the average. Source: Doing Business, World Bank

Table 2 continued

Cost	Cost measures the fees levied on a 20-foot container in U.S. dollars. All the fees associated with completing the procedures to export or import the goods are included. These include costs for documents, administrative fees for customs clearance and technical control, customs broker fees, terminal handling charges, and inland transport. The cost does not include customs tariffs and duties or costs related to ocean transport. Only official costs are recorded. We take average value of the variable over all years for which data are available and use log values of the average. Source: Doing Business, World Bank
------	---

Table 3 Descriptive statistics of all the variables

	Obs.	Mean	SD	Minimum	Maximum
Documents	106	2.75	0.23	2.2	3.22
Population	106	15.88	1.96	10.76	20.98
Income	106	8.04	0.94	5.85	9.75
Trade-to-GDP ratio (log)	106	4.06	0.46	3.01	5.12
Weighted tariff (log)	106	2.16	0.61	0.53	3.31
Latitude	106	0.24	0.16	0.01	0.67
Regulation	106	108.1	45.28	13	183
Common law	106	0.31	0.46	0	1
Socialist law	106	0.19	0.41	0	1
Catholic	106	0.31	0.46	0	1
Muslim	106	0.29	0.46	0	1
Protestant	106	0.13	0.34	0	1
Ethnic	106	0.49	0.24	0.04	0.93
Area	106	11.95	2.28	5.56	16.62
Time	106	4.03	0.46	3.04	5.19
Cost	106	7.93	0.51	6.76	9.34

2.1 Dependent variable

The main dependent variable is a measure of trade facilitation defined as the number of documents required for exports and imports (*Documents*). The data source for the variable is the World Bank's Doing Business project. Average values of the variable for all years for which data are available (2005–2010) were first computed; log values of these averages were then used to arrive at *Documents*. For the full sample, the values of *Documents* range between 2.2 (St. Kitts and Nevis) and 3.2 (Central African Republic), with the mean value equal to 2.75 and SD of 0.23.

Figure 1 pictures the correlation between *Documents* and our measure of country size (total population of the country (log values)) that is formally defined below. It shows that the number of documents required for exports and imports (trade facilitation) tends to increase (decrease) when country size increases. This relationship is economically large and also statistically significant at less than the 1 % level. Figure 2

Table 4 Correlations between the explanatory variables

	Log of documents	Log of GDP	Log of population	Log of trade-to-GDP ratio	Log of weighted tariff	Latitude	Regulation (EODB)	Common law	Socialist law	Catholic	Muslim	Protestant
Log of documents	1											
Log of population	0.307*** (0.001)	1										
Log of GDP	-0.397*** (0.000)	-0.233** (0.016)	1									
Log of trade-to-GDP ratio	-0.153 (0.117)	-0.315*** (0.001)	0.236** (0.015)	1								
Log of weighted tariff	0.025	-0.105	-0.237**	-0.271***	1							
Latitude	0.055 (0.803)	0.098 (0.286)	0.297*** (0.015)	0.148 (0.005)	-0.346*** (0.000)	1						
Regulation (EODB)	0.363*** (0.000)	0.010 (0.310)	-0.532*** (0.000)	-0.178* (0.069)	0.384*** (0.000)	-0.306*** (0.001)	1					
Common law	-0.148 (0.130)	-0.193** (0.048)	-0.007 (0.943)	0.080 (0.416)	0.145 (0.137)	-0.260*** (0.007)	-0.232** (0.017)	1				
Socialist law	0.069 (0.483)	0.077 (0.435)	0.120 (0.220)	0.280*** (0.004)	-0.383*** (0.000)	0.665*** (0.000)	-0.300*** (0.002)	-0.327*** (0.001)	1			
Catholic	-0.080 (0.418)	-0.146 (0.136)	0.190* (0.051)	-0.121 (0.216)	0.005 (0.957)	-0.233** (0.016)	0.069 (0.480)	-0.176* (0.071)	-0.283*** (0.003)	1		

Table 4 continued

	Log of documents	Log of GDP	Log of population	Log of trade-to-GDP ratio	Log of weighted tariff	Latitude	Regulation (EODB)	Common law	Socialist law	Catholic	Muslim	Protestant
Muslim	0.162* (0.096)	0.245** (0.011)	-0.124 (0.204)	-0.110 (0.304)	0.117 (0.233)	0.141 (0.149)	0.159 (0.103)	-0.152 (0.121)	-0.007 (0.940)	-0.432*** (0.000)	1	
Protestant	-0.052 (0.600)	-0.220** (0.024)	-0.031 (0.756)	0.029 (0.767)	-0.039 (0.694)	-0.204** (0.036)	-0.180* (0.065)	0.533*** (0.000)	-0.194** (0.046)	-0.262*** (0.007)	-0.251*** (0.010)	1
Ethnic fraction-alization	0.276*** (0.004)	0.214* (0.027)	-0.445*** (0.000)	-0.192** (0.049)	0.125 (0.201)	-0.290*** (0.003)	0.394*** (0.000)	-0.005 (0.959)	-0.225** (0.021)	-0.128 (0.193)	0.158 (0.107)	0.062 (0.531)

This is a power correlation matrix. Figures in parentheses represent statistical significance of the correlation coefficients

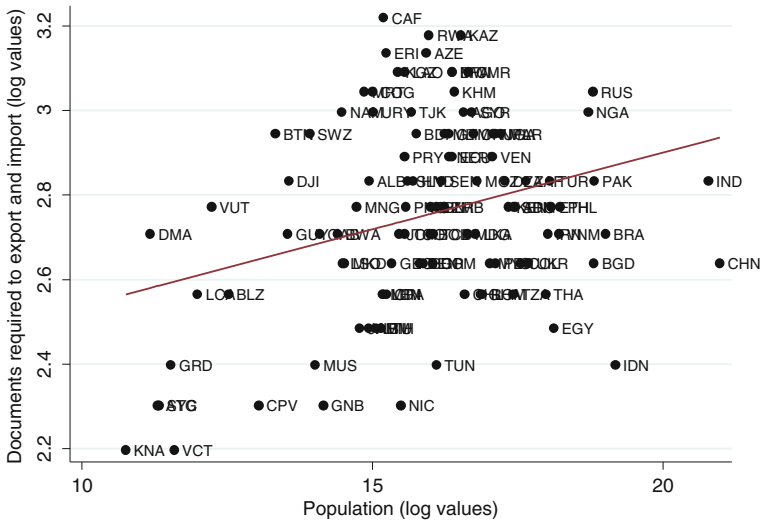


Fig. 1 Correlation between documents to trade and country size. *Note* (1) the horizontal axis plots the value of *Population* and the vertical axis plots the values of *Documents* as defined above. (2) The positive relationship shown is statistically significant at less than the 1 % level. (3) Country abbreviations are taken from World Development Indicators, World Bank

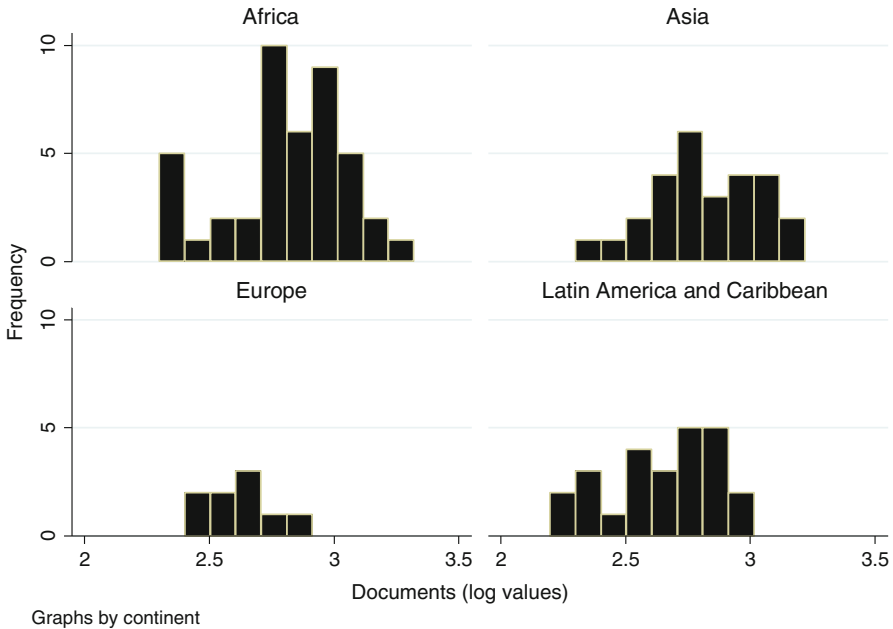


Fig. 2 Average number of documents to trade across continents. *Note* The horizontal axis in each of the graphs plots the values of *Documents* as defined above

shows the full distribution of *Documents* in the sample. The Doing Business data also report on the number of days it takes to clear shipments of exports and imports (*Time*) and the monetary cost of complying with all the procedures involved in exporting

and importing (*Cost*). Like *Documents*, we use the (log of) average values over all years for which data are available (2005 to 2010) for *Time* and *Cost*. We treat *Time* and *Cost* as alternative measures of trade facilitation, although some caution is necessary here since unlike *Documents*, *Time* and *Cost* are outcome measures and hence not trade facilitation measures per se. Further, it is not clear whether *Time* and *Cost* are purely driven by what are commonly considered as trade facilitation measures or by other factors as well. Hence, we treat the regression results for *Cost* and *Time* as purely robustness checks. These results are discussed briefly and contained separately in Sect. 3.3.

2.2 Explanatory variables

Following the literature, we measure country size by the (log of) total population of the country. That is, we first take the average value of total population of a country between 2001 and 2005 and then take the log of the average values to get our main explanatory variable, *Population*. The mean value of *Population* equals 15.9 and the SD is 1.96. Values of *Population* vary between 10.8 (St. Kitts and Nevis) and 21 (China). As mentioned above, existing studies show that population is strongly correlated with macrolevel measures of trade openness (trade-to-GDP ratio, tariff rates), with small countries being more open to trade than large countries. Does a similar relationship hold for trade facilitation as measured by the number of required documents to export and import? This paper attempts to answer this question.

To capture potential nonlinear effects of country size on trade facilitation, we use the square of population as an additional explanatory variable ($Population^2 = Population * Population$). As discussed above, the possibility of a nonlinear relationship between country size and trade facilitation is largely an empirical issue with theory offering very little help in this regard.

The use of lagged values of population and the fact that demographic factors are typically macro in nature compared with our trade facilitation measure that is micro in nature suggests that reverse causality is unlikely to be much of a problem with our estimation. That is, it is highly unlikely that the current level of trade facilitation of the type discussed above could have affected the level of population across countries in the past. However, our regression results could suffer from omitted variable bias problem. To guard against this possibility, we use a number of controls informed by the broader literature on trade openness and country size. As a further check against the omitted variable bias problem, we provide results using the instrumental variable regression method.

One could argue that the richer countries are likely to have fewer numbers of documents (better trade facilitation) than the poorer countries. Also, existing studies have found that income level is typically higher among the relatively small countries. Hence, regression results for the relationship between population and the number of required documents could be spuriously driven if we do not control for income differences across countries. Income level also serves as a broad measure of the overall economic development (quality of institutions, etc.). It is plausible that overall economic development is correlated with the overall quality of trade facilitation or *Documents*. If the

level of overall economic development also varies systematically with country size, the omitted variable bias problem is then evident. To guard against these possibilities, we control for *Income* defined as GDP per capita (PPP adjusted and at constant 2005 USD, log of the average value over 2001–2005).¹ As we show below, controlling for *Income* is important as doing so has a significant impact on the estimated strength of the relationship between *Population* and *Documents*.

Another concern comes from the known relationship between trade openness and country size. That is, small countries are known to be more open to trade than large countries, and greater trade openness could be causally related to better trade facilitation (lower value of *Documents*). This concern again implies an upward bias in our estimate of *Population* if we do not control for trade openness. We check for this possibility by controlling for two measures of trade openness which include exports plus imports as a ratio of GDP (trade-to-GDP ratio, log values) and the overall tariff rate weighted by the volume of bilateral trade for each product (*Weighted tariff*).

We compliment the income measure above with the absolute distance of countries from the equator (*Latitude*). One motivation for controlling for *Latitude* is the same as for GDP per capita since *Latitude* is a proxy for overall development. Another motivation is that geography could be important for trade and hence the quality of trade facilitation. If *Latitude* is also correlated with *Population*, then the omitted variable bias problem is implied.

A natural question to ask here is whether the proposed relationship between *Population* and *Documents* is specific to trade facilitation or if it is part and parcel of a broader relationship between country size and the overall level of regulation. To this end, we control for a measure of the overall level of regulation measured by the World Bank's Ease of Doing Business (EODB) index (*Regulation*). As expected, in our sample, countries with high values of *Regulation* (implying heavier regulation) are also the ones that have significantly higher values of *Documents*. However, our results discussed below show that the relationship between country size and the trade facilitation measure is unique in that it goes well beyond any general relationship between the overall level of regulation and country size.

A number of studies have shown that the legal origin of a country has a significant effect on various aspects of the business climate including the regulation of entry, property, financial development, and the quality of courts (see, for example, [La Porta et al. \(2008\)](#) and [Amin and Haidar \(2012\)](#)). Hence, one might expect some correlation between legal origin and our trade facilitation measure. Although there is little work on how country size varies with the legal origin of countries, if the two happen to be systematically correlated, then the possibility of a spurious correlation with our main results cannot be ruled out. To counter this possibility, we control for the legal origin of countries using dummy variables for the English Common Law and the Socialist Law, with the French Civil Law as the omitted category.

Last, we control for some cultural factors including the main religious group in the country (dummy variables for Catholic, Muslim, and Protestant with the residual

¹ We would like to emphasize here that our income measure (GDP per capita) is a proxy for overall economic development and not for country size. Some studies use income as a proxy for country size, but the income measure they use is an aggregate of total GDP of the country and not GDP per capita.

religious group being the omitted category) and the degree of ethnic fractionalization (*Ethnic*). It is argued that one disadvantage of being large is that large countries are also more diverse. The greater diversity makes it more difficult to closely cater to individual preferences over public goods and in reaching consensus over reforms². Independent of country size, studies have shown that greater ethnic fractionalization has a direct adverse effect on various aspects of overall development and the quality of institutions. Controlling for the main religious group is in the nature of a robustness check, although there is little theoretical or empirical reason to believe that either country size or trade facilitation is strongly correlated with religion.

3 Estimation

3.1 Linear specification

We begin with the results for the linear specification using the OLS estimation method. These results are provided in Table 5, columns 1–8. Without any other controls, the estimated coefficient value of *Population* is positive equaling .036, significant at less than the 1 % level (column 1). In other words, a unit increase in *Population* is associated with an increase in the value of *Documents* by .036. That is, a 1 % point increase in population level is associated with a 3.6 % point increase in the number of documents required to export and import. Alternatively, the estimated coefficient value implies that moving from the smallest to the largest country in terms of *Population* increases the value of *Documents* by .037 or about 36 % of the difference between the highest and the lowest value of *Documents*. This effect is economically large.

Regression results in columns (2)–(5) show that controlling for the income level, trade-to-GDP ratio, or weighted tariff, individually or jointly, does not change our main results. For example, the estimated coefficient value of *Population* remains positive, economically large, and statistically significant at less than the 5 % level even after controlling for GDP per capita and the two trade openness measures (column 5). However, the coefficient value does decline in magnitude from .036 (column 1) to .026 (column 5). This decline is almost entirely due to the control for GDP per capita. While the trade openness measures show very little correlation with the dependent variable, *Income* is strongly negatively correlated with *Documents*.

Given that controlling for income level had a fairly large effect on the estimated coefficient value of *Population*, controlling for *Latitude*, an additional measure of overall development, becomes all the more important. However, column (6) shows that controlling for *Latitude* has little effect on the estimated coefficient value of *Population*, although the significance level of the coefficient declines somewhat to between the 5 and 10 % level (7.2 %). Unlike *Income*, *Latitude* shows a positive correlation with *Documents* implying more required documents among the relatively more developed countries. However, this positive relationship is weak, being statistically insignificant at the 10 % level.

² For more about the role and impact of business regulatory reforms across countries, see Haidar (2009, 2012).

Table 5 Base regression results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable: Documents									
Population	0.036*** [0.003]	0.027** [0.021]	0.0344** [0.012]	0.037*** [0.003]	0.026** [0.045]	0.023* [0.072]	0.021* [0.084]	0.022* [0.091]	0.024* [0.062]
Income		-0.085*** [0.000]			-0.088*** [0.000]	-0.097*** [0.000]	-0.067*** [0.007]	-0.062** [0.023]	-0.059** [0.054]
Trade-to-GDP ratio			-0.032 [0.543]		-0.006 [0.896]	-0.011 [0.815]	-0.015 [0.759]	-0.024 [0.620]	-0.009 [0.866]
Weighted tariff				0.022 [0.561]	-0.015 [0.706]	-0.001 [0.989]	-0.028 [0.494]	-0.027 [0.509]	-0.013 [0.760]
Latitude						0.219 [0.107]	0.257** [0.047]	0.2 [0.188]	0.236 [0.146]
Regulation (EODB)							0.001*** [0.010]	0.002** [0.015]	0.001*** [0.020]
Common law								0.017 [0.740]	0.009 [0.861]
Socialist law								0.047 [0.497]	0.089 [0.276]
Catholic									0.080 [0.248]
Muslim									0.058 [0.382]
Protestant									0.106 [0.198]

Table 5 continued

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Ethnic									0.085 [0.378]
Constant	2.175*** [0.000]	3.012*** [0.000]	2.341*** [0.000]	2.117*** [0.000]	3.107*** [0.000]	3.167*** [0.000]	2.853*** [0.000]	2.829*** [0.000]	2.583*** [0.000]
R ²	0.094	0.206	0.098	0.097	0.207	0.227	0.276	0.279	0.302
Observations	106	106	106	106	106	106	106	106	106

P values in brackets. All regressions use Huber–White robust standard errors. Significance level is denoted by *** (1 % or less), ** (5 % or less) and * (10 % or less)

Next, we control for *Regulation*. Regression results in column (7) reveal that controlling for *Regulation* has almost no effect on either the magnitude or the significance level of the estimated coefficient value of *Population*. However, as predicted, *Regulation* shows a sharp positive correlation with the dependent variable, significant at less than the 1 % level.

Our last set of controls includes legal origin, main religious group, and ethnic fractionalization. Regression results in column 8 show that controlling for legal origin has almost no effect on either the magnitude or significance level of the estimated coefficient of *Population*. Further, regression results also confirm that controlling for the religion and ethnic fractionalization, either jointly or individually, has little effect on the estimated coefficient value of *Population*. For example, with all the above controls in place, adding the dummies for the main religious group and the degree of ethnic fractionalization to the specification only marginally increase the estimated coefficient value of *Population* from .022 (column 8) to .024 (column 9), significant at less than the 10 % level.

One concern with the results discussed above could be statistical significance. That is, while the magnitude of the estimated coefficient of *Population* is not much affected by the various controls except for GDP per capita, its statistical significance level does go down to between 5 and 10 % once we control for *Regulation*, legal origin, religion, and ethnic fractionalization. Does this mean that our results for the *Documents–Population* relationship are somewhat weak? Below, we show that this apparent weakness goes away when we allow for nonlinearity in the *Population–Documents* relationship. Hence, the stated weakness appears to be due to a specification bias. This makes our focus on the nonlinear relationship that much more important.

3.2 Nonlinear specification

As discussed above, we now allow for nonlinearity in the *Documents–Population* relationship.³ We do so by adding the square of population to the various specifications discussed above. The motivation for exploring nonlinearity is already discussed in the previous section.

Regression results for the nonlinear specification are provided in Table 6. Without any other controls, the estimated coefficient value of *Population*² is negative, economically large, and statistically significant at less than the 1 % level. In contrast, the estimated coefficient value of *Population* is positive, economically large, and statistically significant at less than the 1 % level. These results imply that while the relationship between country size and the number of required documents is positive, economically large, and statistically significant at all levels of population in our sample, it is much stronger at low values of *Population* than at high values of *Population*. For example, the estimated effect of *Population* on *Documents* varies between .311

³ We also experimented with a cubic relationship between *Documents* and *Population*. However, we found no evidence of any statistically significant (at the 10 % level or less) cubic relationship. The results for the cubic relationship are available on request from the author. We would like to thank an anonymous referee for pointing out the possibility of a cubic relationship and to check for it.

for the smallest country and .172 for the largest country in our sample. Note that the former is about 1.8 times the latter.

Adding the various controls discussed in Sect. 3.1 does not change the results for the nonlinear specification mentioned in the previous paragraph (columns 2–8, Table 6). The estimated coefficient values of *Population* and *Population*² maintain their respective signs as above and remain significant at less than the 1 % level irrespective of the set of controls (Table 6). Further, the total effect of population (*Population* and *Population*²) on *Documents* is positive and significant at less than the 1 % level in all the specifications in Table 6. As above, the estimated effect of population on the dependent variable is much larger for the relatively smaller countries. More specifically, across the various specifications in Table 6, the estimated effect of population for the smallest country varies between 1.81 and 1.85 times the effects on the largest country. To discuss one example, with all the controls discussed above added to the specification, the estimated coefficient value of *Population* equals .427 and that of *Population*² equals –0.013, both significant at less than the 1 % level (column 8). For a comparison, we restate the corresponding coefficient values without any controls, .457 and –.014 (column 1).

We also experimented with adding squared terms of all the controls discussed above to the final specification in column 9 of Table 6. That is, squared terms for *Income*, trade-to-GDP ratio, weighted tariff, *Latitude*, *Regulation*, and *Ethnic*. The motivation here is to check if the nonlinear effect attributed to *Population* above is actually spuriously driven by the nonlinear effect of the other variables in the specification. However, we found no evidence of this. Controlling for the stated squared terms did not have any qualitative effect on the nonlinear relationship between population and the number of required documents.

3.3 Alternative measures of trade facilitation

We now explore the robustness of our results by using two alternative measures of trade facilitation, *Time* and *Cost*, as the dependent variables. Regression results using these alternative measures were mixed. That is, the country size and trade facilitation relationship mentioned above does continue to hold using *Time* and *Cost* as dependent variables, but only in the nonlinear specification and not in the linear specification. For the linear specification, there is no robust relationship between country size and *Time* and *Cost*. Hence, one could argue that our results for the population and documents relationship cannot be easily generalized to other dimensions of trade facilitation (*Time* and *Cost*, for example). However, a more plausible explanation here is that the specification bias in the linear specification is particularly high for the alternative time and cost measures. Once the specification bias is controlled for, results for all the trade facilitation measures (*Documents*, *Time*, and *Cost*) show similar results. We have also argued above that the results for *Time* and *Cost* should be treated with due caution since these are outcome measures and at best partially driven by trade facilitation measures as understood in the conventional sense. More work is needed to ascertain or reject the generality of our results to alternative measures of trade facilitation in linear and nonlinear specifications.

Table 6 Nonlinear regression results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable: Documents									
Population	0.457*** [0.000]	0.368*** [0.000]	0.476*** [0.000]	0.512*** [0.000]	0.410*** [0.000]	0.394*** [0.000]	0.353*** [0.001]	0.448*** [0.000]	0.427*** [0.001]
Population ²	-0.014*** [0.000]	-0.011*** [0.000]	-0.014*** [0.000]	-0.015*** [0.000]	-0.012*** [0.000]	-0.012*** [0.000]	-0.011*** [0.001]	-0.014*** [0.000]	-0.013*** [0.001]
Income		-0.067*** [0.002]			-0.058** [0.012]	-0.067*** [0.003]	-0.045** [0.049]	-0.022 [0.405]	-0.027 [0.359]
Trade-to-GDP ratio			-0.064 [0.176]		-0.027 [0.553]	-0.03 [0.511]	-0.031 [0.495]	-0.063 [0.190]	-0.051 [0.313]
Weighted tariff			0.063 [0.105]		0.025 [0.556]	0.035 [0.423]	0.009 [0.843]	0.013 [0.765]	0.018 [0.683]
Latitude						0.179 [0.160]	0.214* [0.076]	0.118 [0.381]	0.138 [0.339]
Regulation (EODB)							0.001*** [0.029]	0.002*** [0.008]	0.002*** [0.006]
Common law								0.093* [0.088]	0.085 [0.158]
Socialist law								0.105 [0.113]	0.128 [0.102]
Catholic									0.054 [0.427]
Muslim									0.036 [0.584]

Table 6 continued

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Protestant									0.067
									[0.393]
Ethnic									0.004
									[0.972]
Constant	-1.03	0.245	-0.878	-1.605**	-0.093	0.076	0.144	-0.749	-0.67
	[0.167]	[0.760]	[0.241]	[0.048]	[0.924]	[0.937]	[0.877]	[0.471]	[0.545]
R ²	0.216	0.28	0.23	0.241	0.287	0.3	0.333	0.357	0.364
Observations	106	106	106	106	106	106	106	106	106

P values in brackets. All regressions use Huber–White robust standard errors. Significance level is denoted by *** (1 % or less), ** (5 % or less) and * (10 % or less)

4 Instrumental variables regressions

To increase our confidence against the omitted variable bias problem with our results above for the relationship between population and the number of required documents, we now explore the instrumental variables regression method. This method requires identifying an instrument for *Population* such that while the instrument is well correlated with *Population*, it should not have any direct effect on the dependent variable, *Documents*. We follow Rose (2006) in using the (log of) total land area of the country (*Area*) as the instrument. While it is natural to expect countries with larger area to have a larger population, there is no reason to believe that land area should have any effect on the number of documents required to export and import except through its impact on population. That is, no direct effect of the instrument on the dependent variable is expected.

Table 7 shows the results for the first stage of the IV regression where we regress *Population* on *Area*, with and without the various controls discussed above (columns 1–8, Table 7). For all the specifications in Table 7, *Area* shows a large positive correlation with *Population*, significant at less than the 1 % level. For example, the simple regression of *Population* on *Area* without any other controls yields an R^2 value of 0.677 (column 1). That is, about 67.7 % of the variation in *Population* can be predicted from the variation in *Area*.

We take the predicted or the instrumented values of *Population*, $Population^{IV}$, from Table 7 and use these values in place of *Population* for the corresponding specifications (various controls). We note that the various controls in the IV regressions (for the linear and the nonlinear specification) are treated as included instruments in that they are included in the first as well as the second stage of the IV regressions. Regression results for the linear specification using the instrumented values of *Population* are provided in Table 8. These regressions confirm that the linear relationship between *Population* and *Documents* is indeed positive, economically large, and statistically significant. In fact, unlike in the OLS specification where we found the population–documents relationship to be somewhat weak in some of the specifications (significant between 5 and 10 % level), the IV results in Table 8 show that the estimated coefficient of population is always significant at less than the 1 % level.

We repeat the IV regression exercise for the nonlinear specification. To this end, we take the predicted value of population from the first stage IV regressions in Table 7 (as above) and use these predicted values and their squared values in place of population and population squared, respectively. The resulting second stage IV regression results are provided in Table 9. These results confirm what we found earlier for the OLS specification. That is, the number of required documents increases with population, but this effect is much stronger at relatively low levels of population than at high levels of population. The estimated coefficient values of *Population* and $Population^2$ duly instrumented are positive and negative, respectively, and individually significant at less than the 1 % level.

The IV regression results serve to increase our confidence that the relationship between population and the number of required documents is indeed causal and not driven either by reverse causality or omitted variable bias problem. We believe that the nonlinearity in the relationship highlighted above has important implications for

Table 7 First stage IV regression results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable: Population									
Area	0.708*** [0.000]	0.697*** [0.000]	0.682*** [0.000]	0.710*** [0.000]	0.667*** [0.000]	0.663*** [0.000]	0.688*** [0.000]	0.687*** [0.000]	0.722*** [0.000]
Income		-0.128 [0.227]			-0.095 [0.339]	-0.113 [0.298]	-0.241** [0.047]	-0.232* [0.066]	-0.284** [0.028]
Trade-to-GDP ratio			-0.581** [0.033]		-0.601*** [0.030]	-0.609*** [0.031]	-0.572*** [0.034]	-0.589*** [0.036]	-0.673*** [0.003]
Weighted tariff				0.04 [0.842]	-0.141 [0.477]	-0.114 [0.578]	0.03 [0.895]	0.054 [0.814]	-0.124 [0.507]
Latitude						0.403 [0.552]	0.187 [0.778]	-0.136 [0.864]	-1.053 [0.213]
Regulation (EODB)							-0.007*** [0.030]	-0.007*** [0.028]	-0.006*** [0.021]
Common law								-0.084 [0.781]	0.351 [0.222]
Socialist law								0.168 [0.642]	-0.064 [0.847]
Catholic									-0.474* [0.069]
Muslim									-0.072 [0.783]
Protestant									-1.514*** [0.001]

Table 7 continued

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Ethnic									-1.247** [0.025]
Constant	7.418*** [0.000]	8.580*** [0.000]	10.090*** [0.000]	7.313*** [0.000]	11.418*** [0.000]	11.489*** [0.000]	12.519*** [0.000]	12.575*** [0.000]	14.387*** [0.000]
R ²	0.677	0.681	0.695	0.677	0.698	0.698	0.713	0.714	0.772
Observations	106	106	106	106	106	106	106	106	106

P values in brackets. All regressions use Huber–White robust standard errors. Significance level is denoted by *** (1 % or less), ** (5 % or less) and * (10 % or less)

Table 8 Second stage IV regression results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable: Documents									
Population/IV	0.066*** [0.000]	0.057*** [0.000]	0.067*** [0.000]	0.068*** [0.000]	0.060*** [0.000]	0.058*** [0.000]	0.050*** [0.000]	0.051*** [0.000]	0.049*** [0.000]
Income		-0.070*** [0.001]			-0.072*** [0.001]	-0.079*** [0.000]	-0.054*** [0.032]	-0.047* [0.084]	-0.051* [0.087]
Trade-to-GDP ratio			0.012 [0.830]		0.043 [0.421]	0.039 [0.472]	0.026 [0.603]	0.016 [0.755]	0.028 [0.579]
Weighted tariff				0.032 [0.400]	0.012 [0.759]	0.022 [0.598]	-0.008 [0.842]	-0.009 [0.820]	0.007 [0.862]
Latitude						0.155 [0.258]	0.203 [0.111]	0.159 [0.303]	0.224 [0.151]
Regulation (EODB)							0.001** [0.015]	0.002** [0.018]	0.002** [0.015]
Common law								0.03 [0.562]	0.012 [0.816]
Socialist law								0.045 [0.528]	0.092 [0.244]
Catholic									0.1 [0.117]
Muslim									0.052 [0.412]
Protestant									0.144* [0.064]

Table 8 continued

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Ethnic									0.06 [0.535]
Constant	1.704*** [0.000]	2.409*** [0.000]	1.642*** [0.000]	1.608*** [0.000]	2.169*** [0.000]	2.224*** [0.000]	2.102*** [0.000]	2.053*** [0.000]	1.910*** [0.001]
Observations	106	106	106	106	106	106	106	106	106
Observations	107	107	107	107	107	107	107	107	107

P values in brackets. All regressions use Huber–White robust standard errors. Significance level is denoted by *** (1 % or less), ** (5 % or less) and * (10 % or less). *Population*^{*IV*} is the instrumented value of *Population* taken from the corresponding columns of Table 6

Table 9 Second stage IV regression results with nonlinear specification

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable: Documents									
Population ^{IV}	.530*** [.000]	.367** [.012]	.588*** [.000]	.584*** [.000]	.427*** [.010]	.389** [.016]	.429*** [.004]	.471*** [.001]	.574*** [.000]
Population ^{IV} squared	-.015*** [.002]	-.010** [.037]	-.017*** [.001]	-.017*** [.002]	-.012*** [.029]	-.011** [.044]	-.012*** [.012]	-.014*** [.003]	-.017*** [.000]
Income		-.059*** [.006]			-.051** [.021]	-.059*** [.007]	-.034 [.121]	-.021 [.378]	-.018 [.515]
Trade-to-GDP ratio			-.033 [.542]		.013 [.803]	.012 [.819]	-.003 [0.952]	-.023 [.660]	-.031 [.542]
Weighted tariff				.052 [.149]	.03 [.443]	.038 [.356]	.013 [.756]	.012 [.769]	.036 [.416]
Latitude						.146 [.246]	.182 [.130]	.12 [.393]	.137 [.343]
Regulation (EODB)							.001** [.026]	.002** [.016]	.002*** [.009]
Common law								.057 [.220]	.052 [.306]
Socialist law								.07 [.277]	.123 [.100]
Catholic									.074 [.258]
Muslim									.045 [.479]

Table 9 continued

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Protestant									.129*
									[.061]
Ethnic									-.008
									[.933]
Constant	-1.76	0.002	-2.06**	-2.30*	-0.67	-0.34	-0.81	-1.21	-2.12**
	[.103]	[.999]	[.044]	[.056]	[.598]	[.787]	[.467]	[.256]	[.015]
Observations	106	106	106	106	106	106	106	106	106

P values in brackets. All regressions use Huber–White robust standard errors. Significance level is denoted by *** (1 % or less), ** (5 % or less) and * (10 % or less). *Population*^{IV} is the instrumented value of *Population* taken from the corresponding columns of Table 6. *Population*^{IV} squared is the square of *Population*^{IV} values

the broader literature on country size and overall development, business climate, and the quality of institutions. That is, it is possible that the reason why the literature fails to find any significant effect of country size on various economic variables (other than trade openness) is that existing studies are exclusively focused on the linear relationships. This could lead to serious specification bias if the true relationship is actually nonlinear as in our case.

5 Conclusions

The effect of country size on various economic variables has remained largely elusive to economists. Trade openness is a notable exception with a number of studies showing that smaller countries are more open to trade than the large countries. The present paper extends this finding by showing that in addition to the more conventional macrolevel trade openness measures, such as the trade-to-GDP ratio and tariff rates, small countries perform better than large countries in terms of trade facilitation too. That is, the number of documents required for export and import clearance, a measure of trade facilitation, tends to increase sharply with country size proxied by total population. An additional contribution of the paper lies in showing that the country size and trade facilitation relationship is highly nonlinear—much stronger at relatively low levels of country size. More research is needed to ascertain or reject similar nonlinearity between country size and various other economic variables.

Acknowledgments The authors would like to thank two anonymous referees for helpful comments. Mohammad Yasser Safa provided inspiration. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the Governments they represent.

References

- Alesina A (2002) The size of countries: does it matter? Discussion paper no. 1975. Harvard Institute of Economic Research, Harvard University, Cambridge, MA, USA
- Alesina A, Spolaore E (2003) The size of nations. MIT Press, Cambridge
- Alesina A, Wacziarg R (1998) Openness, country size and government. *J Public Econ* 69(3):305–321
- Amin M, Haidar JI (2012) The cost of registering property: does legal origin matter? *Empir Econ* 42(3):1035–1050
- Djankov S, Freund C, Pham CS (2010) Trading on time. *Rev Econ Stat* 92(1):166–173
- Freund C, Weinhold D (2002) The internet and international trade in services. *Am Econ Rev* 92(2):236–240
- Haidar JI (2009) Investor protections and economic growth. *Econ Lett* 103(1):1–4
- Haidar JI (2012) The impact of business regulatory reforms on economic growth. *J Jpn Int Econ* 26(3):285–307
- Hertel TW, Walmsley T, Itakura K (2001) Dynamic effect of the new age free trade agreement between Japan and Singapore. *J Econ Integr* 16(4):446–484
- Porta L, Rafael FL, Silanes AS (1999) The quality of government. *J Law Econ Org* 15(1):222–279
- Rafael LP, Lopez-de-Silanes F, Shleifer A (2008) The economic consequences of legal origins. *J Econ Lit* 46(2):285–332
- Root H (1999) The importance of being small. Unpublished manuscript
- Rose A (2006) Size really doesn't matter. In search for a national scale effect. *J Jpn Int Econ* 20(4):482–507
- Schwarz P (2007) Does capital mobility reduce the corporate–labor tax ratio? *Public Choice* 130:363–380
- Wilson JS, Mann CL, Otsuki T (2003) Trade facilitation and economic development: a new approach to quantifying the impact. *World Bank Econ Rev* 17(3):367–389