Trade Facilitation and Country Size

Mohammad Amin, Jamal Ibrahim Haidar

Documents de Travail du
Centre d'Economie de la Sorbonne

UNIVERSITÉ PARIS 1 – PANTHÉON - SORBONNE –

CES Working Papers

Trade Facilitation and Country Size

Mohammad AMIN, Jamal Ibrahim HAIDAR

2012.45

Maison des Sciences Économiques, 106-112 boulevard de L'Hôpital, 75647 Paris Cedex 13
http://centredeconomiesorbonne.univ-paris1.fr/bandeau-haut/documents-de-travail/
ISSN : 1955-611X

Trade Facilitation and Country Size

Mohammad Amin* and Jamal Ibrahim Haidar**

May, 2012

It is argued that compared with large countries, small countries rely more on trade and therefore they 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 the 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 non-linear, much stronger for the relatively small than the large countries. Third, contrary to what the 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 GDP than the large countries.

L'activité commerciale des petits pays est souvent considérée comme relativement plus intense, de sorte qu'ils apparaissent plus enclins à adopter des mesures de libéralisation commerciale. Cet article étend cette idée en dépassant les mesures traditionnelles d'ouverture au commerce à travers l'utilisation d'une nouvelle mesure reflétant le degré de facilités commerciales des pays à échanger. Précisément, nous analysons le lien entre la taille des pays et le nombre de document strictement nécessaire pour exporter ou importer. Cette étude met en avant trois résultats. Premièrement, les facilités à échanger sont relativement plus importantes pour les pays de petite taille. Ensuite, la relation existante entre la taille des pays et les facilités commerciales est non-linéaire. Si elle est très marquée pour les petites économies, elle devient moins évidente pour les grand pays. Enfin, contrairement à ce que les études existantes suggèrent, la relation que nous mettons en évidence ne dépend pas de la forte propension des petits pays à exporter.

Keywords: Country size, Trade facilitation, Openness

JEL: F10, F13, F14, F15, F50, F55, H10, K00, L5, O20, 024

*Enterprise Analysis Unit, World Bank, Washington DC, 20433. Email: mamin@worldbank.org.
**Paris School of Economics: Jamal.Haidar@ParisSchoolofEconomics.Eu

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. Smallness 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 micro level, 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 percentage 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 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 non-linearity 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. Implication of this non-linear relationship for the broader literature is 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 2000), time to clear shipments at ports (Djankov et al. 2010) and standards harmonization and automating customs procedure (Herter 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 it needs 2 documents 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.

The motivation for exploring the possibility of non-linearity 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 non-linear or linear. 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.

The plan of the remaining sections is as follows. In section 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 section 3. To raise out confidence against possible endogeneity concerns with our main results, instrumental variable regression results are provided in section 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 possible endogeneity problem with our main estimation results. All regression results use Huber-White robust standard errors and have been checked for possible outliers that may have unduly large effect 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.

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 to 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 standard deviation of 0.23.

Figure 1 pictures the correlation between *Documents* and country size. It shows that the number of documents required for exports and imports (trade facilitation) tends to increase (decrease) when country size increases. Figure 2 shows the full distribution of *Documents* in the sample. The Doing

Business data also reports 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 section 3.3.

2.2 *Explanatory variables*

Following the literature, we measure country size by population. 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 standard deviation 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 macro level 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 non-linear effect 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 non-linear 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 suggest 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). 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. To 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 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 is it 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, financial development and the quality of courts (see, for example, La Porta et al., 2008). 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 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 reforms. Independently 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 the religion.

3. Estimation

3.1 Linear specification

We begin with the results for the linear specification using the Ordinary Least Squares estimation method. These results are provided in Table 5, columns 1 to 8. Without any other controls, the estimated coefficient value of *Population* is positive equaling .036, significant at less than the 1 percent level (column 1). In words, a unit increase in *Population* is associated with an increase in the value of *Documents* by .036. That is, a 1 percentage point increase in population level is associated with a 3.6 percentage 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 percent 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 either 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 percent 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 above (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 5 and 10 percent level (7.2 percent). 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 percent level.

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 percent 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 percent 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 to 10 percent 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 non-linearity in the *Population-Documents* relationship. Hence, the stated weakness appears to be due to a specification bias. This makes our focus on the non-linear relationship that much more important.

3.2 Non-linear specification

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

Regression results for the non-linear 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 percent level. In contrast, the estimated coefficient value of *Population* is positive, economically large and statistically significant at less than the 1 percent 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, at the estimated effect of *Population* on *Documents* varies between .311 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 section 3.1 does not change the results for the non-linear 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 percent 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 percent 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 to 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 percent level (column 8). For a comparison, we restate the corresponding coefficient values without any controls, .457 and -.014, respectively (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 non-linear effect attributed to *Population* above is actually spuriously driven by the non-linear 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 non-linear 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 non-linear 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 non-linear specifications.

Section 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 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 percent level. For example, the simple regression of *Population* on *Area* without any other controls yields an R-squared value of 0.677 (column 1). That is, about 67.7 percent 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 non-linear 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 to 10 percent level), the IV results in Table 8 show that the estimated coefficient of population is always significant at less than the 1 percent level.

We repeat the IV regression exercise for the non-linear 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 increase 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 percent 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 non-linearity in the relationship highlighted above has important implications for 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 because existing studies are exclusively focused on the linear relationships. This could lead to serious specification bias if the true relationship in actually non-linear as in our case.

5. Conclusion

The effect on 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 macro level trade openness measures such as 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 non-linear – much stronger at relatively low levels of country size. More research is needed to ascertain or reject similar non-linearity between country size and various other economic variables.

References

[1] Alesina, Alberto and Romain Wacziarg (1998), "Openness, Country Size and Government," *Journal of Public Economics*, 69(3): 305-321.

[2] Alesina, Alberto (2002), "The Size of Countries: Does It Matter?" Harvard Institute of Economic Research, Discussion Paper No. 1975, Harvard University, Cambridge, Massachusetts.

[3] Alesina, Alberto and E. Spolaore (2003), *The Size of Nations*, MIT Press, Cambridge.

[4] Djankov, Simeon, Caroline Freund and Cong S. Pham (2010), "Trading On Time," *Review of Economics and Statistics*, 92(1): 166-173.

[5] Freund, Caroline and Diana Weinhold (2000), "On the Effect of the Internet on International Trade," International Finance Discussion Papers No. 693, Board of Governors of the Federal Reserve System.

[6] Hertel, Thomas W., Terrie Walmsley and Ken Itakura (2001), "Dynamic Effect of the "New Age" Free Trade Agreement between Japan and Singapore," *Journal of Economic Integration,* 16(4): 446-84.

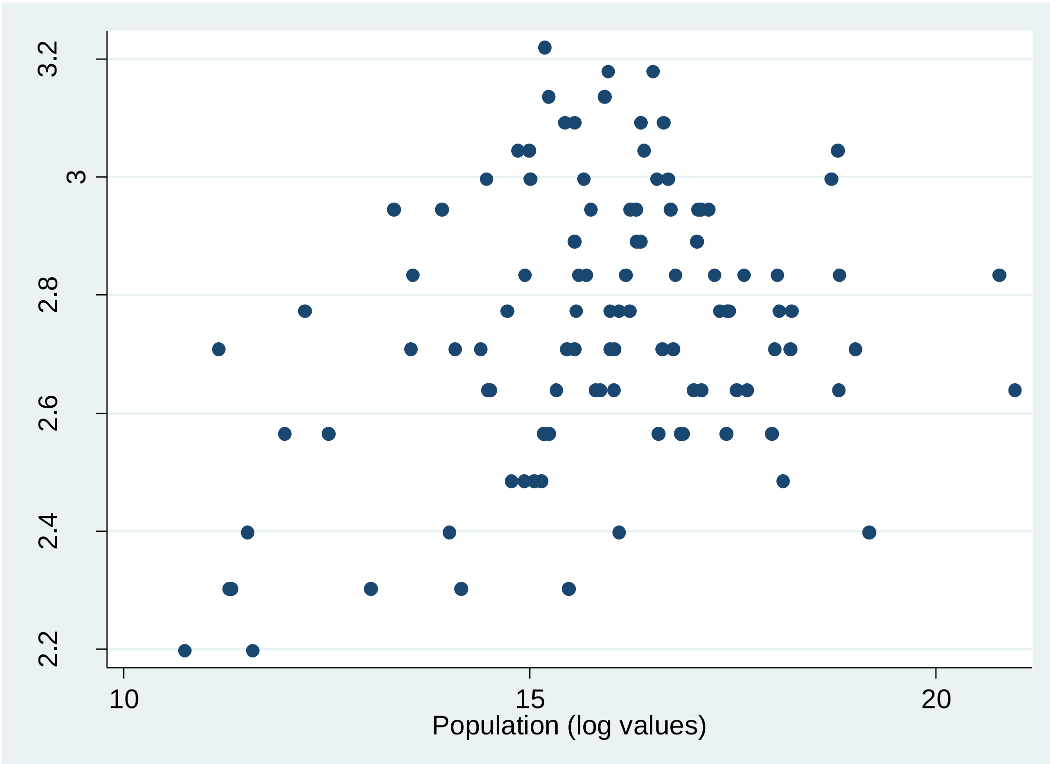
[7] La Porta, Rafael, Florencio Lopez-de Silanes, Andrei Shleifer and Robert Vishny (1999), "The Quality of Government," *Journal of Law, Economics and Organization*, 15(1): 222-79.

[8] La Porta, Rafael, Florencio Lopez-de Silanes and Andrei Shleifer (2008), "The Economic Consequences of Legal Origins," *Journal of Economic Literature*, 46(2), pp. 285-332.

[9] Rose, Andrew (2006), "Size Really Doesn't Matter: In Search for a National Scale Effect," *Journal of the Japanese and International Economies*, 20(4): 482-507.

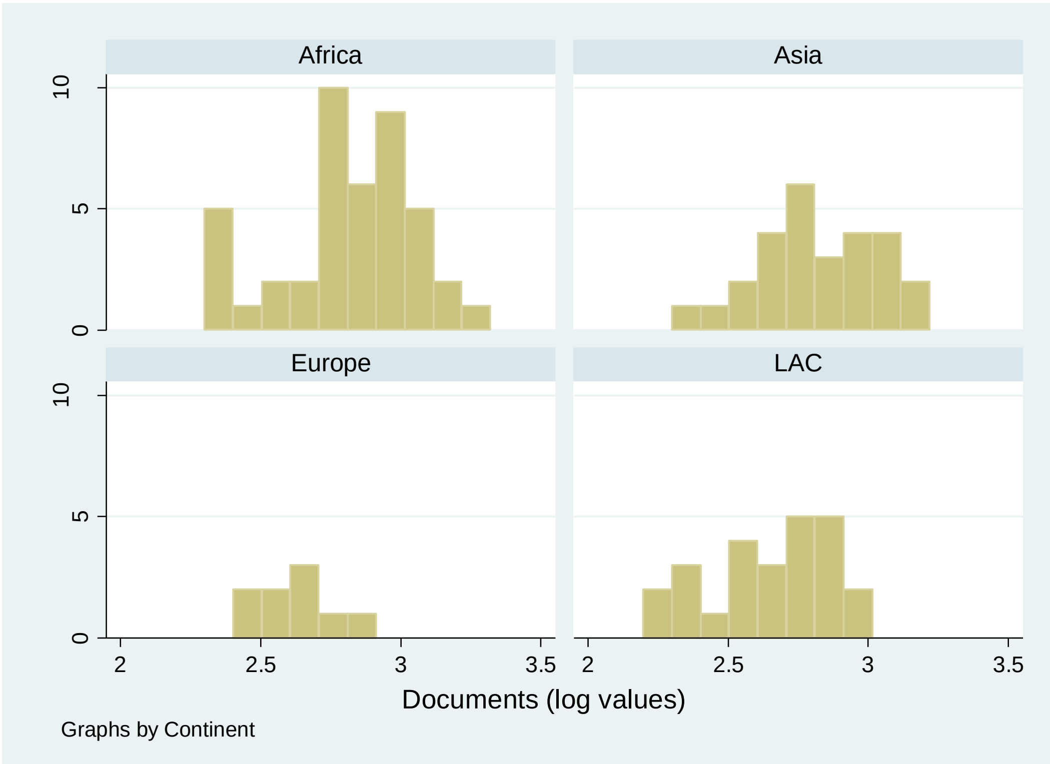
[10] Wilson, John S., Catherine L. Mann and Tsunehiro Otsuki (2003), "Trade Facilitation and Economic Development: Measuring the Impact," World Bank Working Paper No. 2988, World Bank, Washington D.C.

Figure 1: Correlation between Documents to Trade and Country Size



Note: The horizontal axis plots value of *Population* and the vertical axis plots values of *Documents* as defined above.

Figure 2: Average number of documents to trade across continents



Note: The horizontal axis in each of the graphs shown plots the values of *Documents* as defined above.

Table 1: List of countries included in the sample

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

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 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 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 et al. (2003), *Journal of Economic Growth*, June 2003; Table A1.
Area	Log of total land area of the country is 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 applying 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.
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	Std. Dev.	Min	Max
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

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***	1										
	(0.001)											
Log of GDP	-0.397***	-0.233**	1									
	(0.000)	(0.016)										
Log of Trade to GDP ratio	-0.153	-0.315***	0.236**	1								
	(0.117)	(0.001)	(0.015)									
Log of Weighted tariff	0.025	-0.105	-0.237**	-0.271***	1							
	(0.803)	(0.286)	(0.015)	(0.005)								
Latitude	0.055	0.098	0.297***	0.148	-0.346***	1						
	(0.592)	(0.319)	(0.002)	(0.131)	(0.000)							
Regulation (EODB)	0.363***	0.010	-0.532***	-0.178*	0.384***	-0.306***	1					
	(0.000)	(0.310)	(0.000)	(0.069)	(0.000)	(0.001)						
Common Law	-0.148	-0.193**	-0.007	0.080	0.145	-0.260***	-0.232**	1				
	(0.130)	(0.048)	(0.943)	(0.416)	(0.137)	(0.007)	(0.017)					
Socialist Law	0.069	0.077	0.120	0.280***	-0.383***	0.665***	-0.300***	-0.327***	1			
	(0.483)	(0.435)	(0.220)	(0.004)	(0.000)	(0.000)	(0.002)	(0.001)				
Catholic	-0.080	-0.146	0.190*	-0.121	0.005	-0.233**	0.069	-0.176*	-0.283***	1		
	(0.418)	(0.136)	(0.051)	(0.216)	(0.957)	(0.016)	(0.480)	(0.071)	(0.003)			
Muslim	0.162*	0.245**	-0.124	-0.110	0.117	0.141	0.159	-0.152	-0.007	-0.432***	1	
	(0.096)	(0.011)	(0.204)	(0.304)	(0.233)	(0.149)	(0.103)	(0.121)	(0.940)	(0.000)		
Protestant	-0.052	-0.220**	-0.031	0.029	-0.039	-0.204**	-0.180*	0.533***	-0.194**	-0.262***	-0.251***	1
	(0.600)	(0.024)	(0.756)	(0.767)	(0.694)	(0.036)	(0.065)	(0.000)	(0.046)	(0.007)	(0.010)	
Ethnic Fractionalization	0.276***	0.214*	-0.445***	-0.192**	0.125	-0.290***	0.394***	-0.005	-0.225**	-0.128	0.158	0.062
	(0.004)	(0.027)	(0.000)	(0.049)	(0.201)	(0.003)	(0.000)	(0.959)	(0.021)	(0.193)	(0.107)	(0.531)

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

Table 5: Base regression results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable: *Documents*									
Population	0.036***	0.027**	0.034**	0.037***	0.026**	0.023*	0.021*	0.022*	0.024*
	[0.003]	[0.021]	[0.012]	[0.003]	[0.045]	[0.072]	[0.084]	[0.091]	[0.062]
Income		-0.085***			-0.088***	-0.097***	-0.067***	-0.062**	-0.059*
		[0.000]			[0.000]	[0.000]	[0.007]	[0.023]	[0.054]
Trade to GDP ratio			-0.032		-0.006	-0.011	-0.015	-0.024	-0.009
			[0.543]		[0.896]	[0.815]	[0.759]	[0.620]	[0.866]
Weighted tariff				0.022	-0.015	-0.001	-0.028	-0.027	-0.013
				[0.561]	[0.706]	[0.989]	[0.494]	[0.509]	[0.760]
Latitude						0.219	0.257**	0.2	0.236
						[0.107]	[0.047]	[0.188]	[0.146]
Regulation (EODB)							0.001***	0.002**	0.001**
							[0.010]	[0.015]	[0.020]
Common Law								0.017	0.009
								[0.740]	[0.861]
Socialist Law								0.047	0.089
								[0.497]	[0.276]
Catholic									0.080
									[0.248]
Muslim									0.058
									[0.382]
Protestant									0.106
									[0.198]
Ethnic									0.085
									[0.378]
Constant	2.175***	3.012***	2.341***	2.117***	3.107***	3.167***	2.853***	2.829***	2.583***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
R-squared	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).

Table 6: Non-linear regression results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable: *Documents*									
Population	0.457***	0.368***	0.476***	0.512***	0.410***	0.394***	0.353***	0.448***	0.427***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.001]	[0.000]	[0.001]
Population2	-0.014***	-0.011***	-0.014***	-0.015***	-0.012***	-0.012***	-0.011***	-0.014***	-0.013***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.001]	[0.000]	[0.001]
Income		-0.067***			-0.058**	-0.067***	-0.045**	-0.022	-0.027
		[0.002]			[0.012]	[0.003]	[0.049]	[0.405]	[0.359]
Trade to GDP ratio			-0.064		-0.027	-0.03	-0.031	-0.063	-0.051
			[0.176]		[0.553]	[0.511]	[0.495]	[0.190]	[0.313]
Weighted tariff				0.063	0.025	0.035	0.009	0.013	0.018
				[0.105]	[0.556]	[0.423]	[0.843]	[0.765]	[0.683]
Latitude						0.179	0.214*	0.118	0.138
						[0.160]	[0.076]	[0.381]	[0.339]
Regulation (EODB)							0.001**	0.002***	0.002***
							[0.029]	[0.008]	[0.006]
Common Law								0.093*	0.085
								[0.088]	[0.158]
Socialist Law								0.105	0.128
								[0.113]	[0.102]
Catholic									0.054
									[0.427]
Muslim									0.036
									[0.584]
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-squared	106	106	106	106	106	106	106	106	106
Observations	0.216	0.28	0.23	0.241	0.287	0.3	0.333	0.357	0.364

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 7: First stage IV regression results									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable: *Population*									
Area	0.708***	0.697***	0.682***	0.710***	0.667***	0.663***	0.688***	0.687***	0.722***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Income		-0.128			-0.095	-0.113	-0.241**	-0.232*	-0.284**
		[0.227]			[0.339]	[0.298]	[0.047]	[0.066]	[0.028]
Trade to GDP ratio			-0.581**		-0.601**	-0.609**	-0.572**	-0.589**	-0.673***
			[0.033]		[0.030]	[0.031]	[0.034]	[0.036]	[0.003]
Weighted tariff				0.04	-0.141	-0.114	0.03	0.054	-0.124
				[0.842]	[0.477]	[0.578]	[0.895]	[0.814]	[0.507]
Latitude						0.403	0.187	-0.136	-1.053
						[0.552]	[0.778]	[0.864]	[0.213]
Regulation (EODB)							-0.007**	-0.007**	-0.006**
							[0.030]	[0.028]	[0.021]
Common Law								-0.084	0.351
								[0.781]	[0.222]
Socialist Law								0.168	-0.064
								[0.642]	[0.847]
Catholic									-0.474*
									[0.069]
Muslim									-0.072
									[0.783]
Protestant									-1.514***
									[0.001]
Ethnic									-1.247**
									[0.025]
Constant	7.418***	8.580***	10.090***	7.313***	11.418***	11.489***	12.519***	12.575***	14.387***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
R-squared	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***	0.057***	0.067***	0.068***	0.060***	0.058***	0.050***	0.051***	0.049***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Income		-0.070***			-0.072***	-0.079***	-0.054**	-0.047*	-0.051*
		[0.001]			[0.001]	[0.000]	[0.032]	[0.084]	[0.087]
Trade to GDP ratio			0.012		0.043	0.039	0.026	0.016	0.028
			[0.830]		[0.421]	[0.472]	[0.603]	[0.755]	[0.579]
Weighted tariff				0.032	0.012	0.022	-0.008	-0.009	0.007
				[0.400]	[0.759]	[0.598]	[0.842]	[0.820]	[0.862]
Latitude						0.155	0.203	0.159	0.224
						[0.258]	[0.111]	[0.303]	[0.151]
Regulation (EODB)							0.001**	0.002**	0.002**
							[0.015]	[0.018]	[0.015]
Common Law								0.03	0.012
								[0.562]	[0.816]
Socialist Law								0.045	0.092
								[0.528]	[0.244]
Catholic									0.1
									[0.117]
Muslim									0.052
									[0.412]
Protestant									0.144*
									[0.064]
Ethnic									0.06
									[0.535]
Constant	1.704***	2.409***	1.642***	1.608***	2.169***	2.224***	2.102***	2.053***	1.910***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[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 non-linear specification

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