Do managers and experts agree? A comparison of alternative sources of trade facilitation data

Alberto Behar[*]   August 2010   VERY PRELIMINARY: COMMENTS APPRECIATED!

This paper constructs country-level aggregates of trade facilitation measures from firm-level responses in the Enterprise Surveys *and compares them with the* Doing Business *indicators, the* Logistics Performance Index *and the* Enabling Trade Index. *Correlations between the data sources are low even for very specific and similar questions. We also use the* Enterprise Surveys *to distinguish between within-country inter-firm variation and between-country variation, finding that the latter accounts for only a quarter of the total. For the purposes of identifying where reform is needed and estimating the relationship between trade facilitation and exports, these findings raise the issue of which form of variation is more informative and which data source is more reliable.*

1   INTRODUCTION

International trade has grown fast in recent years, helped by the signing of multilateral and other trade agreements, but many countries remain relatively isolated. One reason is that transport costs remain high in many parts of the world. While this is in part due to geography – many countries are landlocked or far from attractive markets – man-made and policy characteristics can help as well. Physical infrastructure like roads, communications and ports have been found to be positively associated with trade flows. As a result, large investments have been made by governments and multilateral institutions to improve trade related infrastructure (Behar & Venables, 2010).

However, policymakers and researchers have recently turned their attention to the institutional and administrative barriers to trade. Reforms aimed at removing this type of barrier are often referred to as trade facilitating reforms. The extent to which such barriers exist can be very important because, given recent investment, infrastructure may no longer be the binding constraint. Furthermore, unlike hard infrastructure, it can be cheaper and easier to implement trade facilitating reforms. A number of studies have concluded that countries with a higher degree of trade facilitation – lower administrative/institutional obstacles to trade – tend to have higher trade flows (Wilson et al, 2005; Clarke et al, 2004).

Much of the discussion is in the context of country-level (or bilateral) characteristics that affect transport costs, but a recent literature has considered the role of firms in international trade. Few firms export from any given country, yet international trade appears to be dominated by a few multinationals (Bernard et al, 2007). It then becomes pertinent to make the firm the unit of analysis rather than the country. Why one country exports more than another remains an important question, but asking why one firm exports and another doesn't or why some countries have more exporting firms than others are also good questions. It's also important to establish whether the answers to all the questions have the same implications for the importance of trade facilitation.

Much empirical work on trade facilitation has made use of macroeconomic data in gravity models. Recent work inspired by Melitz (2003) recognizes that firms are not homogenous and uses this insight to explain various features of international trade observed at both the firm- and country-level. Still using macroeconomic data, gravity models have been modified to be able to distinguish between the effects of trade costs on the proportion of firms exporting from a country as well as the quantity that each firm exports (Helpman et al, 2008).

There are a number of macroeconomic sources of country-level trade-related indicators, including the *Doing Business* indicators, the *Logistics Performance Index* and the *Enabling Trade Index*. These all provide measures of trade facilitation that are candidate regressors in gravity models.[1] These differ in their scope, methodology and coverage but one thing they have in common is that the information is drawn from a number of experts but not the firms who are actually exporting. The *Enterprise Surveys* are firm-level surveys that include questions on trade facilitation and international trade and a handful of studies have used this data for selected regions.[2]

Furthermore, the *Enterprise Surveys* have been consistently conducted across a number of countries for the purpose of cross-country comparability.[3] One objective of this paper is to present aggregate summary trade facilitation statistics based on firm-level responses, so we construct various country-level summary statistics of the firm-level responses.

The second objective is to compare the different data sources, especially answers from the microeconomic data with close analogues available in the macro sources. Focusing on measures that facilitate exports (as opposed to imports), we find that various distinct indicators from the same source are highly correlated, but measures of the same characteristic from different sources have a low correlation. This happens even if we are quite precise about the type of question. For example, the *Enterprise Surveys* and *Doing Business* both have information on the number of days exports take to clear ports and customs yet the correlation is only 0.13.

This is important because the countries identified as being in need of reform can differ depending on the source. From a measurement or econometric point of view, for example gravity models of trade,

---

[1] See for example Djankov et al (2010) for *Doing Business*, Behar et al (2009) for the *LPI* and Lawrence et al (2008) for the *ETI*.

[2] Li & Wilson (2009) do so for Asian firms while Balchin & Edwards (2008) do so with African firms.

[3] However, the World Bank notes that the cross-country comparability characteristics of the *Doing Business* dataset are superior. See at http://www.enterprisesurveys.org/Methodology/Compare.aspx

these indicators are usually interpreted as proxies for some underlying country characteristics. Therefore, we discuss whether the methodologies produce alternative proxies for the same thing or whether the issue being investigated becomes different. It also raises the issue of whether one question is more relevant than another or whether one proxy is more reliable than another.

The third objective is to use the *Enterprise Surveys* to compare the variation occurring within-countries (between-firms) with that occurring between countries. We find that cross-country variation explains only one quarter of the total. This suggests macroeconomic studies are ignoring most of the variation in trade facilitation experience and raises the issue of whether a focus on countries is appropriate or whether the within-country variation is more interesting, useful or relevant for measurement and policy.

The answer to this depends in part on what the reason for the variation is. We therefore discuss four interpretations of the cross-firm variation in trade facilitation experience: (i) known firm-specific random draws from a distribution of 'trade facility', (ii) known firm-specific but endogenous trade facility, (iii) uncertain/stochastic trade facility common to all firms but varying for every shipment and, (iv) noise/measurement error.

Section 2 introduces the various data sources, including their scope, coverage and methodology. This includes the *Enterprise Surveys*, where we also explain our approach to producing country-level summary statistics and comparing them with the macroeconomic sources. Section 3 presents and compares the descriptive statistics from each source, noting that the correlations between various measures are low. In general, we present descriptive measures for the whole world but also illustrate with examples from Central Asian countries. In addition, we decompose the variation into that between countries and that within countries, observing that the latter is much bigger.

Section 4 discusses the findings. It provides alternative interpretations of the within-firm variation, attempts to reconcile the data sources and evaluates their relative merits for estimation and policy. We remain unsure about which source is more appropriate, but hope that these empirical findings raise awareness of the potential importance of the various data sources and how they are generated. Until differences are properly reconciled or understood, empirical work should use more than one source in the interests of robustness.

2   DATA CONSTRUCTION

This section discusses each of the original data sources. The core source of microeonomic data is the *Enterprise Surveys*. The macroeconomic datasets are the *Doing Business* indicators, the *Logistics Performance Index* and the *Enabling Trade Index*.

2.1   Enterprise Survey data description

The World Bank *Enterprise Survey* (*ES*) data is available from http://www.enterprisesurveys.org/. There are two "core" or "comprehensive" data sets, which group countries together with comparable survey questions. One set is for the years 2002 through 2006 and the other set is for the years 2006 through 2009. Because many survey questions are different between the two periods, the two datasets are warehoused separately and we concentrate on the latter period. Therefore, we have access to responses from about 40,000 firms across 87 developing countries, taken in various years from 2006 to 2009. This covers a very broad range of topics but we are particularly interested in answers to a number of quantitative and qualitative questions regarding trade and trade facilitation. The following variables are retained from the Enterprise Surveys: [4]

- indirect exports as % total sales
- direct exports as % total sales
- Average days (over the past 2 years) it took to clear export customs from day of arrival at port
- Maximum days (over the past 2 years) it took to clear export customs from day of arrival at port
- Perception of customs and trade regulations as a constraint to business (index from 0-4)

While we will analyse some of the data at the firm level, a key component of our exercise is to calculate summary statistics at the country level. Therefore, for each of these variables, we calculate the mean, median, standard deviation and interquartile range. The measures of central tendency are in principle comparable to the macro indicators. The dispersion measures can have a variety of uses, including comparisons of within- and between-country variation and, depending on what one believes is generating the dispersion, can be informative about the degree of uncertainty faced by firms.

We summarise the statistics across all firms who responded but prefer to use those summarizing responses from exporters for a number of reasons. First, many trade analyses are by definition conditioning on firms who export. Second, it is hard to interpret some answers from non-exporters. For example, if customs and trade regulations are not a constraint on the business, this can be because the requirements are not onerous or simply because this constraint is never encountered by non-exporters

---

[4] We also construct a variable for total exports as a percentage of sales and a dummy for whether or not the firm was an exporter. For completeness, we have import analogues to the export measures as well as information on waiting times for a licence and whether bribes were used for one.

(or importers). Third, for more objective measures, many non-exporting firms are likely to have no experience of actual processes and are therefore likely to be guessing. Fourth, the structure of the questionnaire does not appear to explicitly instruct enumerators to skip these questions for non-exporters, but response rates are much lower for this group.[5] Four countries do not have enough exporters who answered the questions so we effectively have 83 countries.

Further, we provide the above statistics as sample summary statistics but also try to reflect the population summary statistics by taking account of the survey design. When the surveys are conducted, firms of all sizes and ownership structures were interviewed, but certain industries were concentrated on for cross-country comparability. The sampling methodology for Enterprise Surveys is stratified random sampling with replacement. Stratified random sampling groups all population units into homogenous groups and then selects simple random samples from each group. This includes an over-sampling of firms with over 100 employees (World Bank, 2009a). As a result, population estimates must take account of population weights when calculating the means and must account for both the weights and stratification[6] when calculating the standard deviation.

Because the *Enterprise Surveys* aim to visit most countries every three years, a handful got surveyed twice in the 2006-9 period. Although exploiting time series variation within a panel is a fruitful line of enquiry, our exercise only uses the latest survey for those countries.[7]

2.2   Macroeconomic datasets

2.2.1   Doing Business data description and survey years

The *Doing Business* (*DB*) report is produced annually by the World Bank and International Finance Corporation. The 2010 edition (World Bank, 2009b) includes 183 countries, including developed and developing nations. *DB* surveys are conducted in person with local experts, including lawyers, consultants, accountants, freight forwarders and government officials, to verify the *de jure* requirements for each step of the trading process, including each piece of paperwork, payment, and license necessary

---

[5] Overall, response rates can be low as a result. For example, less than 6,000 firms gave a number when asked about the days it takes to clear customs.

[6] For a handful of countries, there was no obvious stratification variable so we assumed one stratum. For those with multiple strata of which some have only one sampling unit, we treat these as certainty equivalents with scaling based on the variances of the other strata.

[7] The timing of the Enterprise Surveys within years varies. In 2008 and 2009 surveys, enumerators record the time, day, month, and year in which the survey is taken. Unfortunately, for 2005, 2006, and 2007 reports, the enumerators do not record the survey date. Among the 2008 and 2009 surveys, there is no consistent pattern for whether the year of the report and year of the survey correspond. That is, about half of the reports from 2009 had surveys conducted in 2009 and the other half in 2008. Therefore, for 2008 and 2009, we use the year in which the majority of surveys were taken as the year of the survey, whereas for 2005, 2006, and 2007, we use the reporting year.

for a representative business to export or import. When surveying the experts, the firms in question are assumed, *inter alia*, to have more than 60 employees, export at least 10% of their sales, be domestically owned and located in the country's most populous city (World Bank, 2009b).[8]

The World Bank "Trading Across Borders" section, located at http://www.doingbusiness.org/ExploreTopics/TradingAcrossBorders/, reports six main measures of trade facilitation:

- number of documents to export or to import
- days to export or to import
- cost to export or import a standard shipping container in dollars

These six variables are available for all years, but the 2010 edition disaggregates the time component for some countries. The breakdowns are:

- days to clear ports (export or import)
- days spent on document preparation (export or import)
- days spent on in-land transport (export or import)

The breakdown is available for exports and for imports and can be accessed through each country's profile. So, for document preparation, the 2010 edition has both the number of documents and the days taken to process them.

We used the 2010 edition in order to have the time breakdowns but also use data from the edition that corresponds to the year that the Enterprise Survey was conducted in each country. This is done with a one year lag because *Doing Business* reports are typically released in the year preceding the report's label, so that data in the report corresponds to the previous year. For example, since the most recent Enterprise Survey was conducted in Albania in 2007, data from the Doing Business 2008 report was used for all Doing Business variables in Albania. For completeness, we also have the averages from the 2005-9 reports.

2.2.2 LPI

Starting in 2007, the World Bank *Logistics Performance Index* (*LPI*) will be based on surveys conducted every two years. The two editions of the data currently available are discussed in reports named *Connecting to Compete: Trade Logistics in the Global Economy* (Arvis et al, 2007, 2010). The reports and data are available at www.worldbank.org/lpi. The 2010 report includes information for 155 countries, including developed and developing countries.

---

[8] This applies to the "Trading Across Borders" component of the survey. Other components have different firm characteristics.

The *LPI* reports six sub-indexes and an overall index: (1) efficiency of customs clearance, (2) quality of trade and transport related infrastructure, (3) ease of arranging competitively-priced shipments, (4) competence and quality of logistics services, (5) ability to track and trace consignments, and (6) frequency with which shipments reach the consignee within the scheduled or expected delivery time. The overall index is constructed from the 6 dimensions using principle components analysis. While both reports share the same components, the 2010 report offers more detailed breakdowns and an expanded emphasis on internal logistics. As a result, the 2010 edition makes a more explicit distinction between local and international logistics although the international part is largely unchanged in content.

The *LPI* draws from a structured online survey receiving nearly 1,000 responses from logistics professionals who are based in international logistics companies in 130 countries. Ten percent of respondents are located in low-income countries, 45 percent in middle-income countries, and 45 percent in high-income countries. Each respondent is asked to rate 8 overseas markets on logistics performance using a qualitative assessment.[9] The 8 markets are different for each respondent, based on the most important export and import markets of their location country, neighboring countries that facilitate their goods transport to ports, and random selection (Arvis et al, 2010).

2.2.3 ETI

The World Economic Forum's *Enabling Trade Index* (*ETI*) is a meta-index that takes unweighted averages of other indicators and whole indices. Of its roughly 55 components, it includes 15 from its own survey, the Executive Opinion Survey, which it carries out annually to ask CEOs and other top business leaders to rank country capacities from 1 to 7. The other 40 or so components include quantitative and qualitative indicators and indices from publicly available sources: International Trade Centre (13 components), World Bank Logistics Performance Index (5), World Bank Doing Business (6), International Telecommunication Union (4), World Economic Forum Global Competitiveness Index (5), United Nations Conference on Trade and Development (UNCTAD) (2), and International Air Transport Association (IATA) (1).

Normalizing each indicator or index to a 1 to 7 scale to match its Executive Opinion Survey, the ETI creates the unweighted average for each of its four sub-indices and then the overall unweighted average of the sub-indices. The sub-indices are

- market access, which measures the extent to which the policy framework welcomes foreign goods into the country and enables access to foreign markets for domestic exporters
- border administration, which assesses the extent to which the administration at the border facilitates the entry and exit of goods

---

[9] The domestic component of the 2010 LPI and some aspects of the 2007 index were also backed up by quantitative information from respondents.

- infrastructure, taking into account whether the country has transport and communications infrastructure to facilitate the movement of goods within the country and across the border
- business environment, which looks at governance, security and the regulatory environment impacting importers and exporters.

In turn, these are based on nine pillars – the mapping to the indices is not explicitly clear – and these pillars are based on a number of individual questions and components. While many of the trade-related variables include measures from sources we have discussed separately in this paper, which means there is some duplication by construction, trade-related components do also come from other sources. The results have since 2008 been published annually in the *Global Enabling Trade Report*, of which the 2009 edition covers 121 developed and developing countries (Lawrence et al, 2009).[10]

2.3   Comparison of coverage

The following matrix shows the number of countries overlapping between the four data sources:

	Enterprise Surveys 2006-9	Doing Business 2010	LPI 2010	ETI 2009
Enterprise Surveys 2006-09	87	87	76	61
Doing Business 2010	87	183	149	121
LPI 2010	76	149	155	115
ETI 2009	61	121	115	121

The main disparity is due to the fact that the *Enterprise Surveys* do not cover developed countries. *Doing Business* has the most countries and its coverage by and large nests that of the other sources.

The obvious difference between the *Enterprise Surveys* and the others is that actual firms are surveyed about their experiences. In particular, a business owner or top executive is interviewed. While accountants or human resource officers may be interviewed for some sections, there is no indication that a person responsible for logistics or operations is asked. The range of firms is wide although we rely predominantly on summary statistics for exporters only. The answers they give can be both objective and subjective. *DB* asks mostly objective questions of a number of local experts and restricts itself to

---

[10] This includes a 2010 version after our dataset was put together. This latest report is available at
http://www.weforum.org/en/initiatives/gcp/GlobalEnablingTradeReport/index.htm

large exporters in a particular city.[11] *LPI* asks logistics professionals in a number of countries about other countries and is by-and-large a perceptions based index. *ETI* is a composite of other macroeconomic sources, including those discussed here, but also includes information from its own survey of CEOs opinions. We return to these issues in section 4.

2.4   Other macroeconomic data

For completeness, we have added in a number of other variables on trade and macroeconomic characteristics that are regularly incorporated in gravity models and other analyses of trade relationships.

To get the most up to date data possible such that it matches some of our trade facilitation data, we sourced information on GDP and the population from the "Historical Data Files" section of the USDA Economic Research Service's International Macroeconomics datasets web site: http://www.ers.usda.gov/data/macroeconomics/.   GDP is real GDP is billions of 2005 US dollars, as obtained directly from the USDA dataset and the population is the number of people. We have this for the years 2005-9 but also an average over the 2006-9 period matching our *Enterprise Survey* coverage.

We have data on country area and whether or not it is landlocked. We also include dyadic data on distance, of which there are various measures and we take the subset based only on the capital city. We have dyadic information on whether countries share a border or a former common colonizer, and whether they share the same language. This data can be found at http://www.cepii.fr/anglaisgraph/bdd/distances.htm.

Trade statistics taken from the IMF Direction of Trade database. The data are available in US dollars but we deflated these to 2005 dollars using the deflator from USDA Economic Research Service. We have two forms of the data. One takes the average over the 2006-9 period to mitigate measurement error and to account for the fact that many of our data sources are not available annually.[12] The second matches the trade year to that of the *Enterprise Survey* for that country. Recall that we also have some indicators of trade activity derived from the *Enterprise Surveys.*

---

[11] Further useful comparisons between *Doing Business* and the *Enterprise Surveys* are available at
http://www.enterprisesurveys.org/Methodology/Compare.aspx
[12] We use the annual data rather than higher frequency versions.

3 DESCRIPTION AND COMPARISON OF TRADE FACILITATION DATA

This section provides descriptive statistics from each of the sources and then compares them with other sources. Our focus is on trade facilitation and, where there is a distinction, on measures related to exports as opposed to imports. We leave such complementary analyses to future research.

3.1 Macroeconomic data sources

The *Logistics Performance Index* (*LPI*) comes in 2007 and 2010 editions. In the bottom left half of the correlation matrix (Table 2), we can see that all sub-components of the 2007 *LPI* are highly correlated with the overall index (column 1) and with each other. Similarly, row 1 shows the 2010 components are also highly correlated with the overall index and the top right half shows they are still highly correlated with each other, although the shipments measure, which is the ease and expense with which one can arrange the shipments of goods overseas, is less correlated with the others.

		LPI 2010 correlations (top right)						
		Overall	Customs	Infrastructure	Shipment	Logistics	Tracking	Timeliness
LPI 2007 correlations (bottom left)	Overall	.	0.96	0.97	0.85	0.97	0.95	0.91
	Customs	0.97	.	0.95	0.77	0.93	0.88	0.85
	Infrastructure	0.97	0.96	.	0.79	0.96	0.90	0.85
	Shipment	0.96	0.91	0.92	.	0.78	0.78	0.69
	Logistics	0.97	0.93	0.94	0.93	.	0.92	0.86
	Tracking	0.96	0.91	0.92	0.90	0.94	.	0.83
	Timeliness	0.92	0.86	0.86	0.85	0.87	0.88	.

Table 2: Correlations within various components of the Logistics Performance Index. Top right of matrix gives correlation between components for 2010 while bottom left gives correlations for 2007 data.

Further, we report that the correlation between the 2007 and 2010 overall indices is 0.90, which suggests some movement by countries over the period. This also is found for the correlation for the ranks and as well as the Kendall Tau rank correlation statistic, which is 0.70. Further, the indications are that, overall, the performance has improved over time. Across the 118 countries, the mean score has increased from 2.75 (out of 5) to 2.9 and the difference between the two samples is significantly different. Figure 1 shows that the 2010 density estimate lies to the right of the 2007 estimate. These results are consistent with Arvis et al (2010), who also identify individual country improvers.
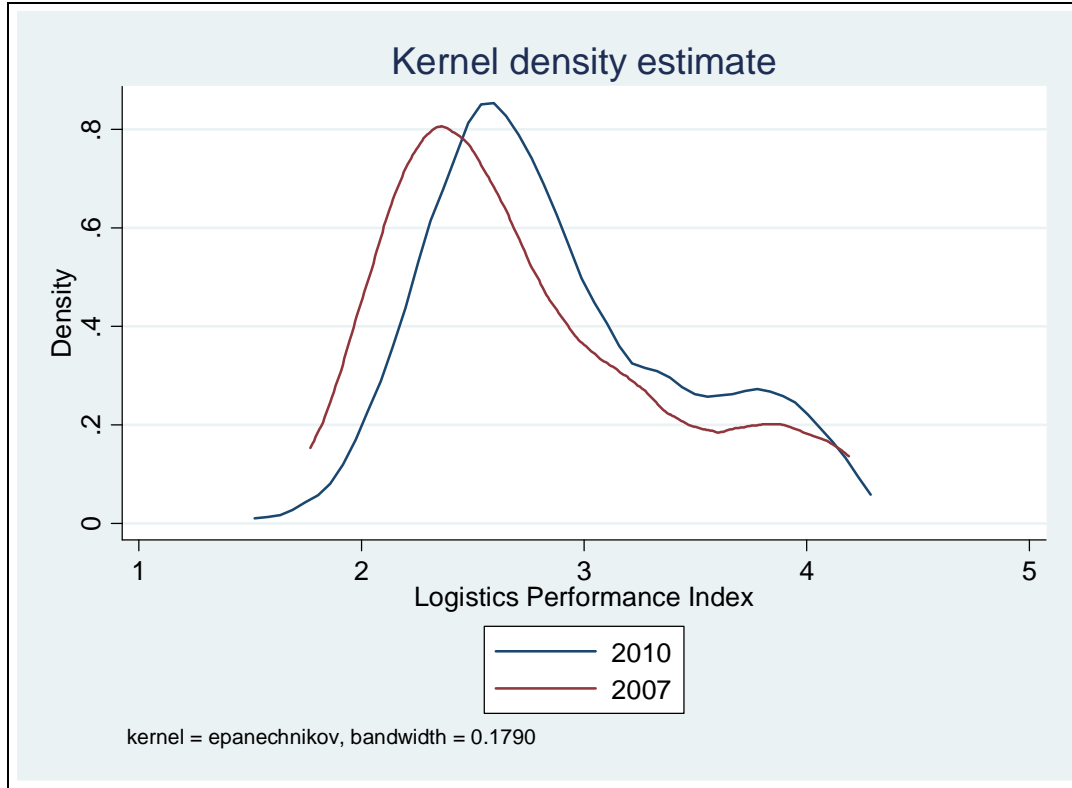
Figure 1: Density estimates of the two editions of the LPI.

Table 3 performs a similar analysis for the *Enabling Trade Index* (*ETI*). The three measures that describe operating conditions – transport infrastructure, business environment and border processes – are highly correlated with the overall index (column 1) and with each other. Market access, which is about tariff and non-tariff trade restrictions, is not closely related to the other measures.

	Overall	Transp. Infra.	Bus. Env.	Border Procs.	Market Access
Overall	1.00				
Transport Infrastructure	0.92	1.00			
Business Environment	0.90	0.81	1.00		
Border Processes	0.96	0.91	0.84	1.00	
Market Access	0.21	-0.09	0.00	0.04	1.00

Table 3: Correlations within ETI components.

Table 4 presents information on the 2010 edition *Doing Business* results.[13] The first row of the table presents the median number of days reported for all the countries. The median is 20 days overall. To have an approximate indication of the breakdown, 7 of these are due to customs and ports delays (further broken down roughly equally) and 12 are due to documentation preparation. We also mention that transit delays last a median of 3 days.

	Total (days)	customs/ ports (days)	ports (days)	customs (days)	documents (days)	documents (number)	cost (dollars)
Median	20	7	4	3	12	6	1190
Total days to export	.						
customs/ ports (days)	0.57	.					
ports (days)	0.44	0.87	.				
customs (days)	0.53	0.78	0.37	.			
documents (days)	0.84	0.40	0.29	0.38	.		
documents (number)	0.56	0.36	0.20	0.43	0.43	.	
cost (dollars)	0.77	0.29	0.20	0.30	0.64	0.37	.

Table 4: median and correlations between Doing Business export trade facilitation measures (all taken from 2010)

The rest of Table 4 presents correlations between various components. The first column suggests that much of the overall correlation in total days is accounted for by documentation delays, while the correlation with the customs/ports component is only 0.57. There is a fairly high correlation between the total number of days and the financial cost of shipping a standard container.[14] The correlation of 0.43 between the number of documents required and the days it takes to complete them is far from perfect.

The purpose of Table 5 is comparison between the three macroeconomic sources. The *DB* measures are lower for better indicators and the others are higher if better. The first panel is for the overall indices. Recall that the *ETI* is in part based on other indices including the *LPI* and *Doing Business*. The first column indicates a fairly high correlation between *ETI* and *LPI* but it is lower with the two *DB* measures.

---

[13] Where there is overlap, the correlation between the 2010 measure and the measure taken from the year in which that country's *Enterprise Survey* was conducted is over 0.95.

[14] For a discussion and comparison of the time and pecuniary costs of shipping goods, see Behar & Venables (2010).

	ETI (overall)	LPI (overall)	DB Cost
ETI (overall)	1.00		
LPI (overall)	0.85	1.00	
DB Cost	-0.51	-0.41	1.00
DB Days	-0.64	-0.56	0.77
	ETI (trans. inf)	LPI (inf)	DB (inland days)
ETI (trans. inf)	1.00		
LPI (inf)	0.90	1.00	
DB (inland days)	-0.37	-0.21	1.00
	ETI (border)	LPI (customs)	DB (customs days)
ETI (border)	1		
LPI (customs)	0.88	1	
DB (customs days)	-0.40	-0.19	1

Table 5: Comparison between macoreconomic sources for overall indices (top panel), transport infrastructure (middle) and border/customs (bottom).

The second panel compares measures of transit and infrastructure, where large correlation between the *ETI* and *LPI* measures suggests the former is made up in large part by the latter. Otherwise, the correlations are fairly low given that we think they are measuring similar things. The third panel concerns customs/border processes. The correlation between *LPI* and *ETI* is again quite high but that between the others is not. In particular, the correlation between the *LPI* perceptions measure and the *DB* objective delays measure is only 0.192.

3.2 Microeconomic data (Enterprise Surveys)

Table 6 presents summary statistics of responses to the question on the degree to which customs procedures and trade regulations are a constraint to the business.[15] Recall we produced variables representing the situation in each country. These are labeled in the columns, while the cross-country summary statistics are labeled in the rows. So, for example, the mean (across countries) of the survey-weighted means is 1.0. More generally, the averages in the first two rows are 1 or just over. The value of 1 corresponds to customs being a minor constraint. So, on average, customs are not perceived as a major constraint. However, for the worst country, the sample median firm gave a value of 3.5, which indicates the constraint is somewhere between major and severe. We will discuss measures of dispersion at a later stage.

---

[15] Unless otherwise indicated, results are based on answers from exporting firms although the answers do not differ materially for the broader set of firms.

		ES Within country summary stats		
		Median	Unweighted Mean	Weighted Mean
Beteween country summary stats	mean	1.3	1.4	1.0
	median	1.0	1.4	1.0
	min	0.0	0.3	0.2
	max	3.5	3.1	2.4

Table 6: Summary statistics on perception of customs and trade regulations as a constraint to business; note 0=no constraint, 1=minor,2=moderate,3=major,4=severe obstacle

Table 7 is analogous to Table 6 but presents an objective measure, namely the response to the question on how many days it takes on average for exported goods to clear customs from the time they arrive at port. The mean across countries of the survey-weighted means is 6.81 days. The median across countries is slightly lower at 5.55 days, which indicates a skewed distribution and/or possible outlier countries. For example, the worst countries report averages of more than 20 days. Further, within countries, the median (in column 1) is typically lower than the means (in the other columns). This may be because many countries have a handful of firms reporting a large value.

		ES Within country summary stats (days)		
		Median	Unweighted Mean	Weighted Mean
Beteween country summary stats	mean	4.05	7.03	6.81
	median	3.00	5.66	5.55
	min	1.00	1.40	1.31
	max	30.00	20.29	20.38

Table 7: Summary statistics on answer to question on the average number of days' for export clearance

Firms were asked about the average number of days to clear customs, but were also asked about the maximum number of days. A comparison between these two questions is presented in Table 8, where the maximum clearance is as expected higher than the average clearance but typically less than twice as high.

		ES Within country average clearance		ES Within country maximum clearance	
		Median	Weighted Mean	Median	Weighted Mean
Beteween country summary stats	mean	3.3	6.8	6.0	10.4
	p50	3.0	5.0	5.3	9.2
	min	1.0	1.0	1.3	2.4
	max	12.5	30.0	16.5	33.2

Table 8: Comparison of answers to questions on average number of days' for export clearance with maximum number of days; note 34 countries

We also note[16] that the between country correlation between the mean responses to average and maximum clearance days is 0.93 but the correlations between these two objective measures and the subjective customs constraint measure were less than 0.3. Further, the correlations between statistics calculated using only exporter responses and those using all firms were high. For the two objective measures, these were in the mid to high nineties. For the perceptions question, some correlations were as low as 0.8, but, as we noted, it is not obvious how to interpret a non-exporter's response to a question on how export procedures constrain the business.

3.3 Comparing macroeconomic and microeconomic sources

This section compares the Enterprise Survey summary statistics with the appropriate macroeconomic sources of data. Table 9 compares the responses to export clearance (from port arrival to customs clearance) with the *DB* data. In principle, the most similar *DB* measure should be the one for clearance of customs and ports. The cross-country median (column 1) is 7 for *Doing Business* and 5.5 from the micro data. The cross-country means of 7.9 and 6.8 are insignificantly different at the 10% level. By these measures, it appears that the two sources are comfortably close.

| | | Cross-country summary stats | | | Pearson correl | | K-Tau correl | |
		median	Mean~	Std. Err.~~	Mean*	Median*	Mean*	Median*
ES	Mean*	5.5	6.8	0.5	.	.	.	.
	Median*	3	4.1	0.5	0.78	.	0.61	.
Doing Business	Total (ES year)	24	30.2	1.9	0.33	0.31	0.19	0.21
	Total (2010)	20	29.1	1.9	0.36	0.36	0.19	0.23
	Ports	4	4.6	0.4	0.14	0.27	-0.06	-0.01
	Customs	3	3.3	0.3	0.06	0.22	0.04	0.14
	Customs/ ports	7	7.9	0.6	0.13	0.30	-0.01	0.08

Table 9: Comparing microeconomic and macroeconomic measures of delays (in days) to clear exports. * refers to question on the *average* days it takes for goods to clear customs from the point of arrival at the port and is either the sample median or the population weighted mean. ~ mean is average across countries and ~~ std. error is of the estimated mean. Total refers to the number of days it takes for goods to clear exports, including all steps, for 2010 (2010) and the year corresponding to the year of the Enterprise Survey for that country (ES year) respectively.

However, the correlations in the right half of the table indicate a different story. In the bottom row, we see that the correlation between the within-country means and the *DB* customs/ports delay is only 0.13. The correlation between the within country medians and the same *Doing Business* customs/ports delay is also low at 0.30. Because one is often concerned about the ranks of the countries, and for robustness,

---

[16] Results available on request

we also present the Kendal Tau rank correlation. By this measure, the correlation is even lower (or negative). In fact, the correlation is insignificant. [17]

There are some higher Pearson correlations in the table, even though they in theory shouldn't be. For example, the 2010 total delay and the within-country means/medians have a correlation of 0.36. While it is not clear how high one might expect these to be, the comparison across answers to different questions seen in the previous subsections yielded much higher correlations. On this basis, the correlations in Table 9 are very low.

While Table 9 compared *Doing Business* with the *ES* responses to average clearance times, Table 10 compares the same *DB* data to the firms' responses on maximum clearance times. The *DB* customs/ports measure (last row) is still higher than the typical within-country median responses (2nd row), which is perhaps surprising given the latter is about maximum delays. However, it is (insignificantly) lower than the typical within-country mean responses (1st row). Although the within-country median responses to this question have a Pearson correlation as high as 0.61 with the *DB* measures, this is not the case for the within-country means or when using the rank correlation.

		Cross-country summary stats			Pearson correl		K-Tau correl	
		median	Mean	Std. Err.	Mean*	Median*	Mean*	Median*
ES		Mean*	9.2	10.4	1.0	.	.	.	.
		Median*	5	6.8	0.9	0.77	.	0.64	.
Doing Business	Total (ES year)	24	28.6	2.1	0.12	0.42	-0.06	0.10
	Total (2010)	20	26.2	2.1	0.19	0.54	-0.07	0.09
	Ports	4	5.1	0.7	0.23	0.61	-0.05	0.07
	Customs	3	3.4	0.2	-0.12	0.10	-0.11	-0.01
	Customs/ ports	7	8.5	0.9	0.16	0.54	-0.12	0.02

Table 10: Comparing microeconomic and macroeconomic measures of delays (in days) to clear exports. * refers to question on the *maximum* days it takes for goods to clear customs from the point of arrival at the port and is either the sample median or the population weighted mean. ~ mean is average across countries and ~~ std. error is of the estimated mean. Total refers to the number of days it takes for goods to clear exports, including all steps, for 2010 (2010) and the year corresponding to the year of the Enterprise Survey for that country (ES year) respectively.

To emphasise the point that the correlations are low and to gain insights into why the rank correlation gives particularly low measures, Figure 2 presents four scatter plots. The x-axis has the same *DB* measure in all cases, but whether using the within-country mean of the average clearance question (top left), the within-country median answer to that question (top right), the within-country mean of the maximum clearance question (bottom left) or the within-country median answer to that question, there is very little evidence of a systematic positive relationship.

What little positive correlation is picked up by the Pearson measure is being driven by a handful of observations but the rank measure is not influenced in this way. The rank measure is not necessarily

---

[17] We also used the Spearman rank correlation and examined measures based on unweighted means, responses from all firms (not just exporters), and so on.

superior; after all, it may be more important that both measures generally identify the countries with extremely high delays. However, this is not done consistently either. In the bottom left panel, the country with a microeconomic measure above 30 and a low *DB* measure is Bolivia. Two firms gave answers in excess of 210 days but the summary statistic is based on a reasonable number (72) of firms. In contrast, Angola has a low *ES* value (10.5) but a *DB* measure above 30 in the top right panel. Here, the answer is based on only 5 responses and one more/less observation could have made the median 30.
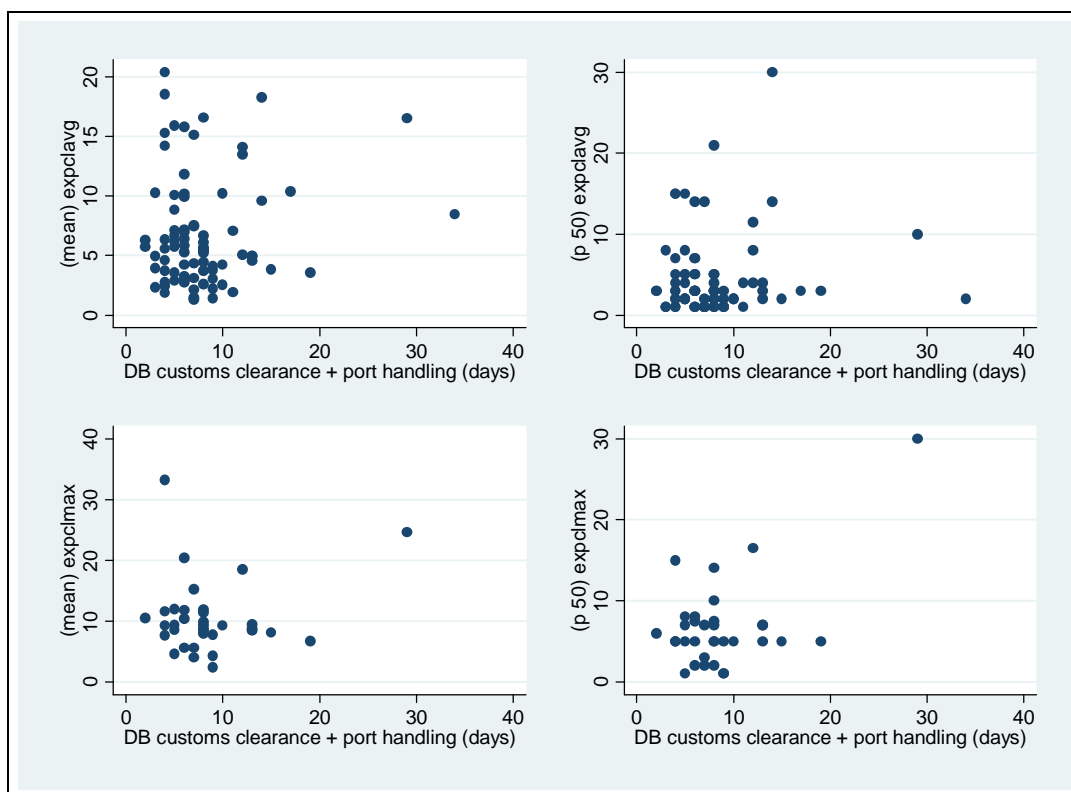

Figure 2: Relationships between macroeconomic and microeconomic measures of customs/ports clearance.

For further comparison, we list and rank all the countries and the mean export clearance average days in the appendix. By this measure, the best performing countries are Botswana and Namibia, who are both part of the Southern African Customs Union. The worst performers are Mongolia and Tajikistan. The appendix also includes the analogous *Doing Business* measures. Azerbaijan, Mongolia, Tajikistan and Mauritius do well by one measure and poorly by the other. Micronesia, Angola, Venezuela, Republic of Congo and Samoa are identified as being poor performers by both sources. The Baltic countries rank well according to both measures.

Having compared objective measures (in days) across sources, Table 11 compares perceptions/index measures in the *Enterprise Surveys, ETI* and the *LPI*. The bottom left presents Pearson correlations while the rank correlations are in the top right. The microeconomic and *ETI* measures are relatively highly

correlated using either measure, especially if benchmarked against the correlation between the *ES* mean and the *ES* median. The correlation with the *LPI* measure is lower. These should not be as comparable with each other as those in the previous two tables, but one might still have expected a higher correlation between measures of similar concepts. Overall, the overall picture for trade facilitation is that the correlations are low.[18]

		K Tau Correlations (top right)			
		ES mean*	ES median*	ETI	LPI
Pearson Correlations (bottom left)	ES mean*	.	0.57	-0.38	-0.38
	ES median*	0.75	.	-0.37	-0.37
	ETI Border	-0.56	-0.60	.	0.53
	LPI Customs	-0.34	-0.32	0.88	.

Table 11: Correlations between microeconomic and macroeconomic index measures of customs/border clearance. * indicates within-country summary statistic of response to question on degree to which customs is a constraint to the business (higher number implies greater constraint)..

3.4 Focus on Central Asia

To revisit some of the comparisons so far and with a view to the next subsection, we present the statistics for individual countries. To compare within a region, we choose those in our dataset that are located in Central Asia. This region is of particular interest to those working on trade facilitation. Its geographical location means it tries to serve the European market but is sufficiently far for speed and cost issues to be important. Further, because many countries are landlocked, border controls are encountered often, so inefficiencies can accumulate.

Table 12 presents sample cross tabulations of firm-level perceptions of the degree to which customs/trade regulations constrain their business as well summary statistics. In total, the mean is 1.42 (bottom right), which places the average Central Asian firm between 'minor' and 'moderate' obstacle. The equivalent measure across all non Central Asian firms is 1.43, so on average, this does not appear to be a generally high constraint for the region.[19]

	Armenia	Azerb.	Belarus	Georgia	Kazakh.	Moldova	Russia	Tajik.	Turkey	Ukraine	Uzbek.	Total
no obstacle %	41	26	31	39	33	22	28	29	51	27	15	39
minor %	11	30	15	15	15	15	16	17	19	20	20	18
moderate %	31	22	23	17	30	23	23	13	14	22	46	20

[18] We briefly checked for other comparable criteria, like the import equivalent to the export measures and measures of transit constraints or infrastructure quality, and found consistently low correlations.
[19] The value of 1.43, which is a mean taken across all firms, is higher than the cross-country mean of the firm-level means reported earlier, which had a value of 1.0.

major %	9	19	25	20	15	18	16	21	8	23	12	14
severe %	7	4	6	9	7	22	18	21	7	8	7	10
Firms #	54	27	65	46	27	60	193	24	583	173	41	1293
mean (0-4 range)	1.24	1.61	1.57	1.43	1.63	2.35	1.92	1.94	1.04	1.83	1.61	1.42
sd (0-4 range)	3.34	3.28	0.70	2.73	1.66	2.25	1.29	4.25	1.30	1.21	2.01	1.37

Table 12: Tabulations of firm perceptions of extent to which customs/trade regulations are an obstacle to their business. The (probability weighted) total mean is across all firms in Central Asia, as is the standard deviation.

Moldova's value of 2.35 places its average firm between 'moderate' and 'major', while Russia and Tajikistan have values of close to 2. The standard deviation also hints at variation across firms within countries. Given the bounded range of the answers, a cross-tabular analysis is perhaps more informative. In Moldova, Russia and Tajikistan, the firms are more or less uniformly distributed across the levels of severity. These three countries distinguish themselves in being the only ones where more than 10% of firms indicated the constraint is severe. Uzbekistan has a roughly symmetric distribution with a clear mode at 'moderate'. Turkey has a skewed distribution in which the proportion of firms decreases at each level of severity. On average, Turkish firms are the least concerned about customs/trade regulations.

Do the macro indicators lead to the same results? Table 13 presents measures from the *Enabling Trade Index* and two editions of the *Logistics Performance Index*. Some indicators are missing but all three measures place the Central Asian countries below the world average. While they are not supposed to be measuring the exact same thing, this is at odds with the firm-level summary. Perhaps the quality is lower but many firms have adapted.

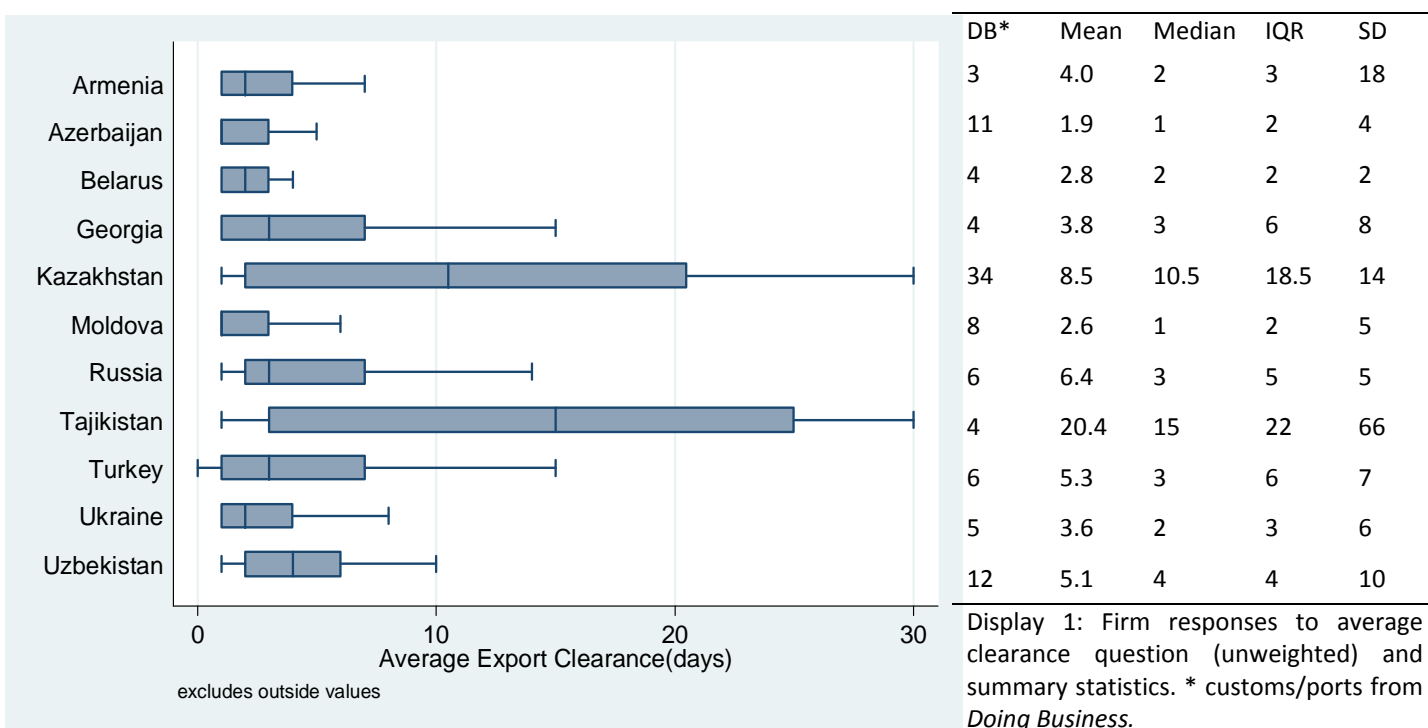|	Armenia	Azerb.	Georgia	Kazakh.	Moldova	Russia	Tajik.	Turkey	Ukraine	Uzbek.	Asia	World
ETI Border	3.25	2.91	.	2.27	3.59	2.82	2.4	4.05	3.07	.	3.05	4.03
LPI Customs '10	2.1	2.14	2.37	2.38	2.11	2.15	1.9	2.82	2.02	2.2	2.22	2.60
LPI Customs '07	2.1	2.23	.	1.91	2.14	1.94	1.91	3	2.22	1.94	2.15	2.55

Table 13: Customs / border quality as measured by ETI (1-7 index) and LPI (1-5 index).

Unlike Table 12, Moldova has a better *ETI* than the Central Asian average and the *LPI* measures put it in the middle. For Russia, *ETI* puts it at third worst but this value is close to that for the middle countries, while the *LPI* measure shows it was among the worst but improved. Tajikistan continues to be the worst or among the worst in Asia, but recall its worldwide position depends on the data source. As for the *Enterprise Survey*, Turkey is the best performer.

We turn our attention to export clearance days. Display 1 presents the answers given by exporting firms to the question on how many days exports take to clear customs together with the summary statistics in

the accompanying table. Most countries have a number of statistical outliers, which may have an impact on the country-level means, but these are not presented in the box plots.[20]

The table identifies Kazakhstan as requiring over a month for goods to clear ports/customs while Uzbekistan and Azerbaijan need almost two weeks. For reference, we note the world-wide median is 7 days. The days are much lower according to the summary statistics of the firm-level responses; the means for these countries are 8.5, 5.1 and 1.9 days respectively. In reverse, the Tajikistan firm responses have a mean of 20 days and a median of 15, while *Doing Business* reports only 4. Both Tajikistan and Kazakhstan have median responses that lie well above the cross-country median of this summary statistic (3 days). These two countries also have mean responses that lie above the cross-country mean of this summary statistic.

	DB*	Mean	Median	IQR	SD
Armenia	3	4.0	2	3	18
Azerbaijan	11	1.9	1	2	4
Belarus	4	2.8	2	2	2
Georgia	4	3.8	3	6	8
Kazakhstan	34	8.5	10.5	18.5	14
Moldova	8	2.6	1	2	5
Russia	6	6.4	3	5	5
Tajikistan	4	20.4	15	22	66
Turkey	6	5.3	3	6	7
Ukraine	5	3.6	2	3	6
Uzbekistan	12	5.1	4	4	10

Display 1: Firm responses to average clearance question (unweighted) and summary statistics. * customs/ports from *Doing Business.*

The box plots indicate substantial variation within Kazakhstan and Tajikistan, which suggests that some firms have very different customs experiences to others. The variation is quite large in a number of countries. The interquartile range and standard deviation quantifies this. For example, Armenia's standard deviation is more than four times its mean while the multiple is less than unity for Belarus. This

---

[20] The width of the rectangular boxes gives the interquartile range, with the left and right sides giving the 25th and 75th percentile of answers. The vertical line inside the box gives the median (in Azerbaijan and Moldova, the median is also the 25th percentile). The length of the lines outside the rectangular boxes is determined by the largest data point that is lower than 1.5 times the interquartile range beyond the 75th percentile (to the right) or by the smallest data point that is higher than 1.5 times the interquartile range beyond the 25th percentile (to the left).

variation within countries motivates a comparison of the sources of variation within countries with the variation between countries.

3.5 Within-country variation vs between-country variation

Table 14 focuses on the dispersion of various measures across countries. In the panel with the 'days' heading, we present two country-level summary statistics based on the *Enterprise Surveys*, namely the sample median and the population-weighted mean across firms. While we have already reported the cross-country means of these summary statistics are 6.8 and 4.1, the cross-country standard deviation of the within-country mean is 4.6. Relative to the cross-country mean, the ratio is 0.7. For the closest corresponding *DB* measure, customs/ports, the standard deviation is similar. Thus, while there is variation across countries, it is relatively small compared to the mean. The interquartile ranges are also similar and narrow.

	Days				Indices			
	Clearance ave.		Doing Business		Customs constraint		ETI	LPI
	Mean	Median	Total	Customs/ports	Mean	Median	Border	Customs
Interquartile range	5.2	3.5	16.0	4.0	0.6	1.0	1.5	0.9
Standard Deviation	4.6	4.3	16.8	5.0	0.4	0.9	1.1	0.6
Minimum	1.3	1.0	5.0	2.0	0.2	0.0	2.0	1.5
Maximum	20.4	30.0	102.0	34.0	2.4	3.5	6.5	4.0
Mean	6.8	4.1	24.4	7.8	1.0	1.3	4.0	2.6
Std. Dev / Mean	0.7	1.1	0.7	0.6	0.4	0.7	0.3	0.2

Table 14: Measures of between country variation in clearance days and perceptions. Note columns refer to variables, eg mean is the country-level summary of the firm-level responses, while rows refer to calculated summary statistics across countries, so standard deviation is the variation between countries. ES customs constraint range is 0-4 while ETI and LPI ranges are 1-7 and 1-5 respectively. Std. Dev / Mean is calculated as ratio of the two rows of summary statistics.

The measures of dispersion for indices are not directly interpretable , but they can be compared to the within-country statistics based exclusively on the *Enterprise* Surveys. Table 15 presents various measures of within-country dispersion in columns and summarises these across the world in the rows. For example, the country with the highest interquartile range has a range of 29 days while the median country has a range of 5 days. We also note that the survey design leads to a large distinction between the weighted and unweighted standard deviations, despite the latter also accounting for stratification.

On average, the within-country standard deviation is substantially higher than the between-country standard deviation. The median country has an unweighted within-country standard deviation that is twice as high as the standard deviation of the median in the previous table. By other measures, the discrepancy is even higher. For example, the mean ratio of the (weighted) standard deviation to the mean is 2.8, which is four times as high as 0.7.

Similarly, for the perceptions responses, the within-country (weighted) standard deviation is about ten times the mean (Table 15) while the analogous between-country ratio is less than one (Table 14). This strongly suggests that within-country variation is bigger than between-country variation.

	Days				Customs constraint perception			
	IQR	SD	SD (weighted)	Ratio	IQR	SD	SD (weighted)	Ratio
Mean	7.4	9.4	19.9	2.8	2.0	1.2	11.1	11.6
Median	5.0	8.2	13.7	2.5	2.0	1.3	9.4	11.1
Minimum	0.0	1.0	1.5	0.3	0.0	0.5	1.6	1.4
Maximum	29.0	33.4	97.7	8.6	4.0	1.8	37.6	37.0

Table 15: Measures of within country variation in clearance days and perceptions. Note columns refer to variables eg SD (weighted) is the measure of the variation across firms within a country after accounting for survey design, while rows refer to summary statistics of these variables across countries, so mean is the mean across countries. Ratio is of SD (weighted) to mean calculated for each country.

For an alternative comparison between the two sources of variation, we ran a number of regressions to see how much of the total variation across firms world-wide can be explained by between country variation. Table 16 presents the results for the average days it takes for exports to clear.

	Constant		Country dummies		Country & year dummies	
	unweighted	weighted	unweighted	weighted	unweighted	weighted
Root mean square error	11.186	10.702	10.634	9.305	10.634	9.3053
R^2	0	0	0.1101	0.2556	0.1101	0.2556
p value (countries)	.	.	0.000	0.000	0.000	0.000
p value (years)	.	.	.	.	0.020	0.000

Table 16: Selected statistics from firm-level regressions of export clearance days.

For reference, the first two columns present the root mean square error from pooled regressions of all firm responses on a constant (unweighted and population weighted). Our main indicator of the importance of between-country variation is the R^2, which explains how much variation is explained by country dummies. The R^2 of 0.11 suggests that between-country variation explains a small part of the sample variation. Adjusting this for population weights raises the statistic to 0.2556. The country dummies are jointly significant. In the final two specifications, we see that dummies included to represent the different years in which the survey was conducted are also jointly significant. However, they do not make any meaningful contribution to explaining the variation across firms.[21] Therefore, three quarters of the firm variation remains unexplained. This is consistent with the comparison of standard deviations given earlier and implies that the within-country variation is more substantial than between-country variation.

---

[21] Although significant, the coefficients did not indicate a trend relative to 2006. The macroeconomic data did produce evidence of improvement over time.

Returning briefly to table 15, we note that there is variation across countries in the extent to which there is within-country variation. The country with the lowest dispersion in export clearance days has a ratio of 0.3 while the maximum is 8.6.

4   DISCUSSION

Section 2 discussed and compared the construction of the various data sources including the firm-level responses from the *Enterprise Surveys* and the country-level information from *Doing Business*, the *Logistics Performance Index* and the *Enabling Trade Index*. Focusing on objective and perception or index measures of trade facilitation, section 3 presented summary statistics from these sources and section revealed that within-country variation across firms is bigger than variation between countries. Further, the correlation between sources is quite low – even for what appear to be responses to very similar objective questions. The purpose of this section is to offer potential interpretations and explanations for these related findings as well as their implications.

4.1   Sources of variation: interpretation and usefulness

With a view to expanding a country's international trade, policy-makers are typically interested in reforming country-wide trade facilitation (and broader institutions). However, we have seen the variation in firm experience within countries is large. Further, some authors have demonstrated that a relative low number of firms export (Bernard et al, 2007; Rankin, Soderbom & Teal, 2006).

Gravity models of trade typically focus on cross-country or bilateral variation in trade and trade facilitation.[22] Some disaggregate between products but the question in mind is still typically at the country-level. For example, Djankov, Freund & Pham (2010) find that country-level time delays (taken from *Doing* Business) affect the ratio of time-sensitive to time-insensitive exports. A few use firm- or even firm-product-destination data within a country but exploit the cross-importing country variation experienced by each firm (Bernard et al, 2009; Lawless, 2008,9) to say something about macroeconomic phenomena. Others examine across-firm, within-country variation in trade and trade facilitation. Examples include Dollar et al (2006), Balchin & Edwards (2008) and Li & Wilson (2009).

Conceptually, cross-firm variation is not necessarily the result of firm-specific characteristics. For example, that fact that one firm is close to the coast while another is far is clearly an attribute of the firm. However, the difference in experience between firms can be small if the roads to the coast are

---

[22] Those studying trade facilitation include Clark et al (2004), Hoekman & Nicita (2009) and Wilson, Mann & Otsuki (2005). Most forms are found to have a positive effect on trade.

good or if the main source of delay is getting off the dock onto a ship rather than getting to the dock. So, reforms at the country level can have an impact on differences between firms.

Melitz (2003) builds a model where firms vary in the efficiency with which they produce things such that only some are sufficiently productive to cover the additional costs of shipping goods overseas, so only some export. This cost is the same for all firms, but a reduction in this cost means some firms subsequently find it profitable to export. This introduces a distinction between the intensive and extensive margin of trade, where the former refers to firms exporting more and the latter refers to more firms becoming exporters.[23] Based on this framework, gravity models with macroeconomic data have been used to show that improved logistics quality is associated with an increase along both margins (Behar et al, 2009).

This framework does not explicitly accommodate variation in international transport costs across firms. While the large variation observed within the *Enterprise Surveys* can be interesting, it is by no means clear that the source is econometrically useful. Papers exploiting cross-firm legislation are explicitly or implicitly assuming that the ease with which they export goods is the result of an exogenous random draw. Many would find this assumption difficult to accept.[24] After all, firms may choose to locate close to the coast because they want to trade internationally and the experience of exporting can make them better at dealing with the additional procedures required. As noted, the literature inspired by Melitz (2003) has firms draw their productivity randomly from a distribution. If this applies to the efficiency with which firms make things, why not the efficiency with which they move them? Leaving this potential inconsistency aside, it is important for both the policy question and for the measurement exercises conducted to inform that question that we understand what the variation actually means. We offer some potentially complementary interpretations.

One interpretation, which we have already touched upon, is that the answers given in the firm-level surveys are due to different firm-specific draws from the same distribution. Putting it crudely, firms who drew too low a transport cost productivity do not find it sufficiently profitable to export while those who drew a high productivity do. By this interpretation, variations between firms in their export levels and their trade facilitation responses would be used to measure the extent to which easier export clearance would raise exports. Of course, if the variation in trade facilitation is genuinely random, then policy cannot help some firms be like others. However, national policy can affect the moments of the distribution such that more firms can have sufficiently favourable draws. Certain types of firms could benefit through help; speculative examples include information packs on how to deal with customs or incentives to relocate.

---

[23] To be more precise, this refers to bilateral exports; that is, exports to a new destination. Additional margins can operate within firms. For example, firms can expand the range of products they export to a destination as well as the quantity of each product exported there (see Bernard et al, 2009).
[24] Country-level variation in trade facilitation is potentially endogenous to trade flows too. On the one hand, more trade raises congestion. On the other, some trade infrastructure projects and reforms are only worth investing in at already high trade levels. Nonetheless, the within-country endogeneity problem is arguably more serious than for macroeconomic data because firm characteristics are more malleable. For example, they can choose their location more easily within a country than the country itself.

A second interpretation is that the differences are firm-specific but almost entirely endogenous to observable and unobservable firm characteristics. A refinement has firms draw from different distributions depending on their features. Some may be observable, for example age, size or the education level of the manager. Further within-country analysis could investigate to what extent variation in answers to trade facilitation correlates with other observables and thus produce conditional distribution analogues to the unconditional means and standard deviations described here.

To the extent that some characteristics remain unobserved and that these characteristics are also related to export success, econometric estimates using within-country variation can be compromised. This endogeneity problem has been noted earlier. Furthermore, without knowing what sources of variation across firms are under the policy-maker's control, it is not clear what steps he could take to improve firms' trade facilitation experiences.

A third interpretation is that the variation within countries reflects the stochastic nature of the process. An extreme form says that a draw from the distribution take place every time a set of goods is shipped, that it is not specific to firms, and that the variation across firms merely reflects a limited number of recent experiences. This source of variation is not very useful for measuring relationships between firm-level responses and firm-level trade outcomes.

However, the interpretation alerts one to the possibility that there is uncertainty faced by firms. This manifests itself as ex post uncertainty – how long will the shipment take this time? – but the first interpretation produces uncertainty because firms may only discover their firm-specific draw after they have attempted to export.

If we view within-country variation as an indicator of the uncertainty faced by firms in that country, then the variation is of direct policy relevance. For example, Freund & Rocha (2010) find that African exports are more responsive to transit delays than other sources of delay. Even though transit makes a small contribution to the total on average, they argue this is because of the unpredictability of this contribution. More generally, it is entirely plausible to expect uncertainty to have a negative impact on exports. Uncertainty in delivery times requires delivery to be earlier, which uses up working capital of both customer and seller, while early arrival of goods can place burdens on the buyer's storage space.

To examine uncertainty more systematically, one could compare the relationship between each country's standard deviation of response times and its exports by adding this measure to a gravity specification.[25] An arguably more direct measure of within-firm uncertainty is the gap or ratio between the answer to the question on average export clearance days and the answer to the question on maximum export clearance days. Under certain exogeneity assumptions, this analysis could be conducted across firms within countries as well as between countries.

---

[25] As suggested, not all this variation is uncertainty faced ex post by firms. One may then wish to more accurately capture the uncertainty faced by firms by calculating the residual standard deviation (or standard error) after conditioning on a number of firm-level covariates.

If it's true that uncertainty negatively affects exports, then steps could be taken to reduce it. Further investigation might reveal the uncertainty is to do with random departures from scheduled opening hours at the border post, which has obvious remedies, or to periodic road closures due to heavy rainfall, which can be mitigated with tarred roads.

A fourth interpretation is that the variation across firms is noise or error. In other words, the accuracy given by respondents is doubtful. While this might be uncertainty in the sense discussed above, we are in this case referring to the respondent giving genuinely inaccurate answers of their experience and/or not even knowing what their distribution looks like.[26] Insufficient knowledge may be of direct relevance for export policy and may reveal ignorance to the prospect of exporting. From an econometric viewpoint, measurement error attenuates estimates of any genuine within-country relationship that might exist between trade facilitation and exports. It can also affect the reliability of the summary statistic for the country.

The next section includes a discussion of how the source of the variation affects the reliability of the country-level summary statistic and hence the correlations with the macro data sources.

4.2   Why do the sources have a low correlation and is one superior?

Our comparison between the macroeconomic and microeconomic measures revealed some low cross-country correlations between data sources. In part, this reflects the nature of the question. For example, conditions may be bad in an objective sense, which is partly captured by indices like the *LPI* and its components, but firms may have coped such that this does not affect their business, as noted in the *ES* perceptions measure. Ironically, if prohibitive export processes cause a firm to focus on the domestic market, then it could well say such processes are no constraint to the current operations of the firm.

Similarly, the low correlations could reflect the fact that different issues are being investigated, but this is not entirely satisfactory for at least two reasons. First, while this may provide grounds for legitimate variation, it is not useful when interpreting econometric results. For example, one typically includes an index as an explanatory variable to proxy the truth about how easy it is to export. Various indices are supposed to be alternative proxies of this truth. At a minimum, if is important to see if these underlying proxies give the same message (or at least the same econometric results). Further, these alternative proxies could be combined with common factor analysis to provide a potentially more accurate proxy.[27]

Similarly, the econometrician may include the number of documents required to clear exports as an explanatory variable but should not necessarily interpret this literally as the effect of document numbers on exports. A significant export documentation variable may have an obvious policy

---

[26] Purists may prefer the use of the label "risk" as opposed to "uncertainty" for the third interpretation.

[27] The *LPI* uses principal component analysis to summarise the variation of various distinct but correlated components. The *ETI* aggregates across various measures to create potentially more accurate proxies but uses simple averages rather than letting the implicit weightings be determined statistically.

implication – reduce the number of documents – but this may not be legitimate if it is approximating the general ease with which one can export in the econometrics. Moreover, these literal interpretations can be dangerous, with countries becoming liable to "reforming to the test". That is, reducing the number of documents required and advertising this on *CNN* while still leaving the underlying environment unchanged.[28]

Second, correlations between separate components of the same data source – for example infrastructure and customs indices within the *LPI* – were high. Tellingly, we were able to compare two sources of answers to a precise question, namely average (or maximum) days it takes for goods to clear customs from arrival at port from the *Enterprise Surveys* and the days it takes for goods to clear ports and customs from *Doing Business*. These correlations were still low and certainly lower than those between conceptually distinct components. The latter may be too high due to some form of halo effect experienced by respondents, but the low correlations in absolute terms demand further interrogation.

Earlier, we noted that *Doing* Business is restricted to the largest city and large firms exporting more than 10% of exports. Although the *Enterprise* survey is broader in geographical coverage and we include all exporters in our summary statistics, the sector coverage was limited for survey design reasons. To the extent that sectors and geographical heterogeneity varies across countries, this may introduce additional variation in the country-level summary statistics. Further, as noted in the World Bank's own comparison between sources,[29] the *Enterprise Surveys* are supposed to yield answers that are representative of experiences of actual firms in that country, while *Doing Business* uses a case study to investigate a hypothetical firm in a theoretical situation. This affects the inference drawn from the data. For example, one country taking more days than another could reflect the fact that its goods are more complicated to move or inspect. Therefore, some of the low correlation could be attributable to differences across countries in the direction in and extent to which actual experience varies from hypothetical.

For the firm-level data, senior firm executives were interviewed. For *Doing Business*, a number of local experts including lawyers in the country were asked. Although the case studies permitted the hypothetical firm to avail itself of any means of speeding up the process and assumed it did so, responses are to the *de jure* trade facilitation situation. However, individual firms may know "short cuts", for example bribes, that can make *de facto* delays shorter. To the extent that these short cuts are reflected in the managers' responses, this explains the generally higher country-level averages found in *Doing Business* than in the *Enterprise Surveys*. Furthermore, the low correlation can be due to differences across countries in the availability of short cuts.

With varying degrees of precision, we have entertained the possibility that the issue, context or person asked affects the actual question and hence answer. Putting it differently, the different data sources are drawing from different distributions (or, loosely speaking, different parts of the distribution). In contrast, they could be drawing from the same distribution but summarizing it with more or less reliability. In

---

[28] Moving from significant indices in a regression has the opposite problem because it is not clear what the actual policy response should be.

[29] Found at http://www.enterprisesurveys.org/Methodology/Compare.aspx

other words, who you ask can affect the quality of the answer. *De jure* answers might be unreliable guides to d*e facto* experiences. Moreover, logistics professionals may know more about logistics than CEOs. However, the CEO is talking about his own firm's actual experience while the logistics professional is evaluating a number of foreign countries.

The firm-level data includes some implausibly large values, which does not allay fears that the within-firm variation is due to noise and therefore unreliable. The Law of large Numbers implies that, given enough firm responses, the mean will converge on the true expected value. So, even if the variation is driven by ignorance (or uncertainty or a few lucky/unlucky instances), the average for the country will be accurate if many firms respond. For some countries, the number of respondents to trade related questions is small.[30] While this need not imply a systematic discrepancy in the cross-country mean between sources, it might account for the low correlation between them.

Low numbers are by no means exclusive to selected *Enterprise Survey* data points. The macroeconomic sources strive to consult as many people as possible, but the data is still based on the responses of relatively few people. Each person's response is implicitly an aggregation of a few or many experiences (but recall *Doing Business* is based on a case study). However, we generally do not know whether the answer for a country would be much different if a different set of experts had been asked.[31] Therefore, especially when the *Enterprise Survey* yields responses from enough firms, it's hard to know which data source is more reliable or appropriate.

It is therefore imperative to check for robustness of results to different data sources. To the extent that there are differences in the nature of the question and who answers it, this is important for interpreting the answers. To the extent that this is due to unknown differences or reliability issues, this is important from a pure statistical / data quality perspective.

---

[30] Earlier, we mentioned Angola, which had only 5 responses to a question and where there was a large discrepancy between the data sources.

[31] The *Logistics Performance Index* explicitly notes the stochastic nature of the responses and takes standard errors into account (Arvis et al, 2010).

REFERENCES

Arvis, J, M Mustra, J Panzer, L Ojala & T Naula (2007), 'Connecting to Compete: Trade logistics in the global economy', The World Bank

Arvis, J, M Mustra, L Ojala, B Shepherd & D Saslavsky (2010), 'Connecting to Compete: Trade logistics in the global economy', The World Bank

Balchin, N & L Edwards (2008), 'TRADE-RELATED BUSINESS CLIMATE AND MANUFACTURING EXPORT PERFORMANCE IN AFRICA: A Firm-level Analysis', Journal of Development Perspectives

Behar, A, P Manners & B Nelson (2009), 'Exports and Logistics', Oxford Department of Economics Discussion Paper

Behar & Venables (2010), 'Transport Costs and International Trade', Oxford Department of Economics Discussion Paper

Bernard, A, J Jensen, S Redding & P Schott (2007), 'Firms in International Trade', Journal of Economic Perspectives, 21 (3), 105-130

Bernard, A, J Jensen, S Redding & P Schott (2009), 'The Margins of International Trade: Long Version'

Clarke, X, D Dollar & A Micco (2004), 'Port efficiency, maritime transport costs, and bilateral trade', Journal of Development Economics, 75, 417– 450

Djankov, S, C Freund & C Pham (2010), 'Trading on Time', Review of Economics and Statistics

Dollar, D, M Hallward-Driemeier & T Mengistae (2006), 'Investment Climate and International Integration', World Development

Freund, C & N Rocha (2010), 'What Constrains Africa's Exports?', World Bank Policy Research Working Paper

Helpman, E, M Melitz & Y Rubinstein (2008), 'Estimating Trade Flows: Trading Partners and Trading Volumes', Quarterly Journal of Economics, 123 (2), 441-487

Hoekman, B & A Nicita (2008), 'Trade Policy, Trade Costs, and Developing Country Trade', World Bank Policy Research Working Paper

Lawless, M (2008), Lawless, Martina (2008). 'Deconstructing Gravity: Trade Costs and Extensive and Intensive Margins', CBFSAI Technical Paper

Lawless, M (2009), 'Destinations of Irish Exports: A Gravity Model Approach', Central Bank and Financial Services Authority of Ireland Research Technical Paper

Lawrence, Blanke, Hanouz & Moavenzadeh (2008), 'The Global Enabling Trade Report 2008', World Economic Forum

Lawrence, Hanouz & Moavenzadeh (2009), 'The Global Enabling Trade Report 2009', World Economic Forum

Li & Wilson (2009), 'Trade Facilitation and Expanding the Benefits of Trade: Evidence from Firm Level Data', ARTNeT Working Paper

Melitz (2003), 'The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity', Econometrica

Rankin, N, M Söderbom & F Teal (2006), 'Exporting from Manufacturing Firms in Sub-Saharan Africa', Journal of African Economies

Wilson, J, C Mann & T Otsuki (2005), 'Assessing the Benefits of Trade Facilitation: A Global Perspective', The World Economy, 841-871

World Bank (2009a), 'ENTERPRISE SURVEY AND INDICATOR SURVEYS SAMPLING METHODOLGY', August 29th, 2009

World Bank (2009b), 'Doing Business 2010: Reforming Through Difficult Times'

APPENDIX: CUSTOMS/PORT CLEARANCE BY COUNTRY

This table presents the probability weighted means of responses to the *Enterprise Survey* question on how many days exports take to clear customs from the point of arrival at port from best to worst. It also presents the *Doing Business* statistics and the country ranking.

ES rank	Country	ES days	DB days	DB rank	ES rank	Country	ES days	DB days	DB rank
1	Botswana	1	7	41	43	GuineaBissau	6	8	49
2	Namibia	1	9	61	44	Nepal	6	8	49
3	Bosnia	1	7	41	45	Peru	6	8	49
4	Latvia	2	4	7	46	Panama	6	2	1
5	Azerbaijan	2	11	69	47	Czech Rep.	6	5	19
6	Serbia	2	7	41	48	Chile	6	6	28
7	Bhutan	2	9	61	49	Sierra Leone	6	8	49
8	Estonia	2	3	3	50	Honduras	6	5	19
9	Romania	2	4	7	51	Poland	6	2	1
10	Lithuania	3	4	7	52	Argentina	6	4	7
11	Niger	3	10	66	53	Russia	6	6	28
12	Moldova	3	8	49	54	Rwanda	7	8	49
13	Slovakia	3	6	28	55	Togo	7	5	19
14	Belarus	3	4	7	56	Ecuador	7	6	28
15	Albania	3	5	19	57	Benin	7	11	69
16	FYMOR	3	6	28	58	Colombia	7	5	19
17	Croatia	3	9	61	59	BurkinaFaso	7	6	28
18	ElSalvador	3	7	41	60	Ghana	7	7	41
19	Uruguay	3	6	28	61	LaoPDR	8	7	41
20	Ukraine	4	5	19	62	Kazakhstan	8	34	83
21	DRC	4	19	81	63	Senegal	9	5	19
22	Swaziland	4	8	49	64	Eritrea	10	14	77
23	Georgia	4	4	7	65	Malawi	10	6	28
24	Gabon	4	9	61	66	Philippines	10	5	19
25	Mauritania	4	15	79	67	Mozambique	10	6	28
26	Armenia	4	3	3	68	CapeVerde	10	10	66
27	Guinea	4	9	61	69	Mauritius	10	3	3
28	Bulgaria	4	6	28	70	Samoa	10	17	80
29	Uganda	4	10	66	71	Chad	12	6	28
30	Hungary	4	7	41	72	Congo	14	12	71
31	Burundi	4	8	49	73	Venezuela	14	12	71
32	SouthAfrica	5	13	74	74	Madagascar	14	4	7
33	Guatemala	5	4	7	75	Cameroon	15	7	41
34	Gambia	5	13	74	76	Bolivia	15	4	7
35	Slovenia	5	3	3	77	Kyrgyz Rep.	16	6	28
36	Nicaragua	5	13	74	78	Brazil	16	5	19
37	Uzbekistan	5	12	71	79	Angola	16	29	82
38	Tanzania	5	8	49	80	Cote d'Ivoire	17	8	49
39	Turkey	5	6	28	81	Micronesia	18	14	77
40	Lesotho	5	8	49	82	Mongolia	19	4	7
41	Paraguay	5	8	49	83	Tajikistan	20	4	7
42	Mexico	6	4	7					