Internet Adoption And Firm Exports: A New Instrumental Variable Approach

Jonathan Timmis
University of Nottingham
jonathan.timmis@nottingham.ac.uk

Abstract

This paper investigates how Internet technology affects a firm's choice of export mode, namely, whether a firm exports directly or via intermediaries. This paper contributes to the literature through a novel instrumental variable approach. I match IP addresses to firm locations to construct an exogenous measure of local Internet infrastructure. I find that developing country firms with Internet technology have greater exports than those that do not. Internet access magnifies direct trade and, if anything, may diminish indirect trade. Importantly, these results are robust to consideration of endogeneity.

1. Introduction

"A manufacturing start-up in Boston can connect with a previously impossible-to-reach supplier in China; a marketing agency in New York can instantaneously collaborate with a client in London; a services firm in France can augment its team by having software developed in India." BBC News, 2012

Only a fraction of firms export directly to foreign markets (Bernard et al, 2009). The majority of firms choose not to export or to export indirectly through an intermediary. Recent research has begun to uncover the importance of intermediaries in cross-border trade. Wholesalers comprise 22 per cent of Chinese exports, 17 per cent of Turkish exports and 41 per cent of Chilean imports (Ahn et al, 2011, Abel-Koch, 2011, Blum et al, 2010). Such firms are commonly assumed to facilitate the matching of buyers and sellers across borders (Ahn et al, 2011). I investigate whether Internet technology provides an alternative mechanism.

The Internet is commonly cited as one of the major drivers of globalisation. Internet technology can reduce the cost of matching buyers and sellers, through greater information provision and lower communication costs. The costs of searching for firms with an online presence, such as a website, are typically lower than in the offline market. Email has reduced international communication costs, being quicker and cheaper than long-distance phone calls or postal services. E-commerce allows firms to transact directly with customers, potentially removing links from the supply chain and reducing costs. Indeed, Internet access forms a government policy objective the world over[1].

Surprisingly, there is a paucity of research in this area. Employing country-level measures of Internet penetration, existing studies typically find a positive and significant trade-creating role of the Internet (Freund and Weinhold, 2004, Clarke and Wallsten, 2006, Tang, 2006, Clarke, 2008). However, the use of country-level data renders it difficult to make causal statements due to the plethora of omitted aggregate factors. Existing research on intermediaries is largely descriptive, mostly examining the characteristics of indirect exporters and their destination markets (Ahn et al, 2011, Bernard et al, 2011). This paper provides a first step to illuminating how technology affects a firm's choice of export mode.

---

[1] For example, EU "Digital Agenda for Europe" (EU, 2013), UNESCO "Broadband Commission for Digital Development" (UNESCO, 2013), Kenya's "Vision 2030" (Government of Kenya, 2013).

[2] Accessible from: http://www.enterprisesurveys.org/

I employ a novel instrumental variable approach to developing country firm-level data. I use within-country geographic variation in IP addresses to identify the causal effect of Internet adoption on firm exports. This new instrument measures Internet infrastructure at a remarkably disaggregated level (a lower bound of 10 square miles) and is available virtually worldwide. I combine the instrument with firm location information to provide an exogenous predictor of Internet adoption. The instrument is of particular value for low and middle-income countries, where alternative measures of regional Internet infrastructure are often non-existent.

I find that firms with a website, or those that use email, have greater exports than those that do not. I also identify interesting export composition effects. Firms with Internet access export more directly, whereas there is little effect for indirect exports. My instrument performs well as a predictor of Internet use. Firms that are located in cities with better Internet infrastructure are more likely to have a website or use email. Importantly, these results are robust to consideration of endogeneity. Internet access magnifies direct trade and, if anything, may diminish indirect trade.

Several papers have employed country-level data to measure the effect of the Internet on international trade. Using a panel of mainly developed countries, Freund and Weinhold (2004) find (lagged) growth of Internet users is significantly correlated with export growth. In an earlier paper, Freund and Weinhold (2002) find US imports of services grew faster from countries with greater Internet penetration. Clarke and Wallsten (2006) instrument for Internet proliferation with measures of competition in the telecommunication sector. Using a cross-section, they find Internet penetration promotes exports from developing to developed countries, but not other trade flows.

The literature employing firm-level data is particularly sparse. Clarke (2008) and Ricci and Trionfetti (2012) use the same data source as this paper, a cross-section of World Bank firm-level data from Eastern Europe and Asia. Ricci and Trionfetti (2012) find that exporters are more likely to have email or a website. Clarke (2008) employs the same country-level instrument as Clarke and Wallsten (2006) and finds a strong correlation between Internet access and firm exports. Abramovsky and Griffith (2006) take an instrumental variable approach similar to this paper. They identify a UK firm's ICT investment using within-country geographic variation in household Internet access. They find that UK firms with greater information communication technology (ICT) investment, tend to purchase more offshore services (Abramovsky and Griffith (2006)).

The remainder of the paper is structured as follows. In Section 2 I discuss the firm-level data. In Section 3 I discuss the application and construction of my instrumental variable. In Section 4 I present the econometric specification. I present the results in Section 5, with robustness analysis in Section 6. In Section 7 I conclude.

2. Data

I use firm-level data from the World Bank Enterprise Surveys[2]. The main advantage of this dataset is that it is collected across a wide-range of countries and, crucially, contains information on Internet adoption and the direct and indirect exports of the firm. Indirect exports refer to the use of traders or agents within the firm's home country, who then export the good abroad. So this measure captures the use of domestic intermediaries.

The panel aspect of the World Bank data is rather limited, with the majority of firms included for one year. My instrument is only available from 2005, which constrains firms to being observed at most twice. Table 1 shows the countries and years of my data. These countries have been selected because they include firms located in at least five different city locations; this is discussed further in the next section.

The World Bank employs stratified random sampling to try and achieve a representative sample of firms. Firms are randomly sampled within groups according to firm size, sector and geographic region within a country.

Table 1: Sample of Countries

Afghanistan	2009	Moldova	2005, 2009
Albania	2005, 2009	Morocco	2007
Argentina	2006, 2010	Nicaragua	2006
Brazil	2009	Pakistan	2007
Chile	2006, 2010	Panama	2006, 2010
Ecuador	2006, 2010	Paraguay	2006, 2010
El Salvador	2006, 2010	Peru	2006, 2010
Guatemala	2006, 2010	Turkey	2008
Honduras	2006	Uruguay	2006, 2010
Mexico	2006, 2010	Zambia	2007

[2] Accessible from: http://www.enterprisesurveys.org/

Previous research has typically measured Internet access as a homogenous variable. The World Bank data allows me to disentangle this into two separate measures: whether the firm has a website and whether it communicates with suppliers or customers via email[3]. Nearly half the firms in the sample have a website and just over three-quarters use email, as noted in Table 2.

Although the majority of firms are observed only once, firms are asked to report their sales and labour employment three years prior to the survey. I calculate lagged labour productivity and convert this to US dollars for cross-country comparability. There is a reasonable degree of missing data, with lagged employment and sales being the most problematic variables. These fields are missing for around a third of firms. In cases of missing data, the observations are dropped from the analysis.

As one would expect, only a minority of firms trade (26 per cent) and exports account for a small proportion of overall sales (11 per cent). The majority of these trading firms are direct exporters. A fifth of all firms export directly, with 8 per cent engaging in intermediated trade (and correspondingly 2 per cent engaging in both).

Table 2: Summary Statistics

Variable	Obs.	Mean	Std. Dev.
Percentage Exports / Sales	15,518	0.107	0.257
Proportion of Exporting Firms	15,518	0.255	
Percentage Direct Exports / Sales	16,293	0.084	0.235
Proportion of Direct Exporting Firms	16,293	0.200	
Percentage Indirect Exports / Sales	16,290	0.027	0.130
Proportion of Indirect Exporting Firms	16,290	0.084	
Website	16,007	0.485	0.500
Email	16,316	0.776	0.417
Firm Age (Years)	16,212	21.203	18.389
Lagged Number of Employees	14,753	85.241	293.002
Lagged Log USD Productivity	11,744	9.289	3.031

[3] These are both recorded as binary variables.

3. Instrumental Variable

The recent availability of firm-level data has highlighted the importance of heterogeneous firm characteristics in determining trade flows. Firms are not equally likely to adopt the Internet nor are they equally likely to engage in international trade. Firms that are larger, younger, more skilled-labour-intensive and export-intensive are more likely to adopt information technologies (Haller and Siedschlag, 2011). These same characteristics are also associated with exporting firms. Exporters are typically larger, more productive and skilled-labour-intensive (Bernard et al, 2007; Wagner, 2007). Therefore a key challenge is in disentangling the direction of causality between Internet adoption and exports. The problem would tend to bias the OLS coefficients upwards, overstating the effect of Internet access.

One would also expect measurement error in the data. The World Bank Enterprise Surveys suffer from a non-negligible degree of missing data. In addition, developing country data is generally acknowledged to be less reliable than those in the developed world. Measurement error would bias the OLS coefficients downwards, understating the effect of Internet access.

Instrumental variables provide a mechanism to obtain consistent estimates of the causal effect in the presence of such endogeneity. The instrumental variable needs to satisfy two conditions. Firstly, it must be correlated with the endogenous Internet variable. Secondly, it must be uncorrelated with any other variable of interest. So the effect of the instrumental variable on trade can only be through Internet adoption.

The existing trade literature has largely neglected the endogeneity problem. Rare exceptions include Clark and Wallsten (2006) and Clarke (2008) who both use country-level measures of telecommunications regulations to predict Internet adoption at the country and firm level respectively. However, finding valid country-level instruments is problematic, as it is difficult to imagine a variable whose only affect on trade is through Internet access. Most similar to this paper, Abromovsky and Griffith (2006) use geographic variation in UK household Internet access as an instrument for a firm's ICT investment in that locality. Household Internet access serves as a measure of local Internet infrastructure that is exogenous to the firm.

I employ a novel measure of Internet infrastructure, namely the density of IP addresses (scaled by population), and use within-country geographic variation of this measure to predict firm-level Internet adoption. A significant advantage of this variable is the degree of spatial disaggregation, allowing evaluation at fine levels of geography within a country (a lower bound of 10 square miles). I combine the instrument with firm location information to provide an exogenous predictor of Internet adoption.

Every computer or device connected to the Internet has a unique IP address. I use the density of IP addresses (scaled by population) as a measure of local Internet infrastructure. I posit that firms are more likely to use the Internet in locations with better Internet infrastructure. It seems unlikely that Internet infrastructure would impact firm exports through another channel other than the Internet.

I employ "geo-location" data, which maps IP addresses to physical locations, kindly provided at a discounted rate by MaxMind Inc. The database covers virtually all IP addresses in the world on a monthly basis from January 2005 onwards. The data covers all of the IP version 4 space, with the exception of IP addresses that are not in use or reserved for private networks (such as home networks)[4]. The data are consequently particularly large, for example, containing 3.8 billion IP addresses in December 2010.

As one would expect the mapping to locations is not error-free. For the countries in my sample, the location is correctly assigned to within 25miles for between 64 per cent and 90 per cent of IP addresses (Maxmind, 2013). The locations are corrected and updated on a monthly basis, which generates some noise in the monthly data. I use the average of the monthly data over the year to minimise this noise.

For each IP address the database provides the city name and longitude and latitude coordinates associated with it. MaxMind provide an additional geographic database containing all possible city locations, whether or not they have an IP address associated with them. I use the two files to aggregate the number of IP addresses at all possible locations. The term "city" is used generously in the MaxMind data, with the location corresponding to far more granular locations such as villages, towns and boroughs of a city. For example, within my sample of countries there are around 87,000 "city" locations. There are a number of different city locations that have near identical coordinates. To make the data more manageable, I aggregate such cases to within a two decimal place combination of longitude and latitude (less than 1 square mile).

The city locations are not all symmetric in size either geographically or in terms of population. In order to control for size differences, I weight the number of IP addresses by

---

[4] The data does not contain any IP version 6 addresses. IP version 4 addresses were exhausted in February 2011, with a progression to IP version 6 thereafter. This does not cause a problem for my sample period, which ends in 2010.

the population of that village, town or city borough. Population data is obtained from the closest year of census data from that country's statistical agency. Where necessary, I aggregate city locations to be consistent with the boundaries within the census data.

I match the IP infrastructure variable to each firm using the firm's location information in the World Bank Enterprise Surveys. The quality of this information is very mixed. Argentinian firms provide details of their neighbourhood within Buenos Aires. Conversely, Polish firms do not provide any city information, only their regional location within Poland. I restrict the sample to countries that have firms in at least five different city locations. Therefore countries such as Poland are omitted from the sample.

Typically the mapping is reasonably straightforward. Where the firm location information is less granular than the IP data, I aggregate IP data (and population) to be consistent. Table 3 shows the number of different "city" locations for each country in my sample. For a later robustness check I restrict the analysis to those countries with at least 20 different locations.

Table 3: Number of Different Matched Firm Locations

Afghanistan	5	Moldova	14
Albania	8	Morocco	16
Argentina	68	Nicaragua	35
Brazil	136	Pakistan	13
Chile	64	Panama	18
Ecuador	15	Paraguay	20
El Salvador	37	Peru	8
Guatemala	33	Turkey	14
Honduras	20	Uruguay	21
Mexico	50	Zambia	5

Figure 1 shows the dispersion of Guatemalan firm locations and Figure 2 the locations of Internet infrastructure. The first thing to note is the spread of firm locations within Guatemala. Secondly, the infrastructure data maps closely to the actual firm location. In general, the IP data is more geographically disaggregated than firm location information.

Figure 3 and Figure 4 show the locations of Chilean firms and Internet infrastructure. Unlike Guatemala, Chilean firms are concentrated within a handful of regions of Chile. Firms record both the city and neighbourhood within which they reside. For example, within Santiago this enables me to map firms of each neighbourhood to a separate Internet infrastructure measure (Figure 5 and Figure 6). In general, the instrument is available at a remarkably disaggregated level; within Santiago I am able to map firms to an area of less than 10 square miles.

I construct a second instrument in order to employ over-identification tests for the joint validity of my instruments. I expect IP infrastructure to be a poorer predictor of Internet access for the very largest firms. In mapping firm locations to IP infrastructure I assume a single plant location, whereas the largest firms are likely to have multiple sites. The very largest firms may be less reliant on local Internet infrastructure to determine the timing of Internet adoption, instead being able to afford Internet access via satellite or pay for installation of a dedicated Internet connection. I construct dummy variables for firms that are above the 90^{th} percentile within their country in terms of lagged employees or lagged sales. I interact each of these dummies with IP infrastructure to generate two additional instruments. I use lagged values so these dummies are exogenous to Internet adoption.

Figure 1: Map of Guatemalan Firm Locations (33 different locations)

Figure 2: Map of Guatemalan IP Infrastructure Locations (364 different locations)

Map data ©2013 Google

Figure 3: Map of Chilean Firm Locations (64 different locations)

Figure 4: Map of Chilean IP Infrastructure Locations (606 different locations)

Map data ©2013 Google

Figure 5: Map of Chilean Firm Locations – Santiago City

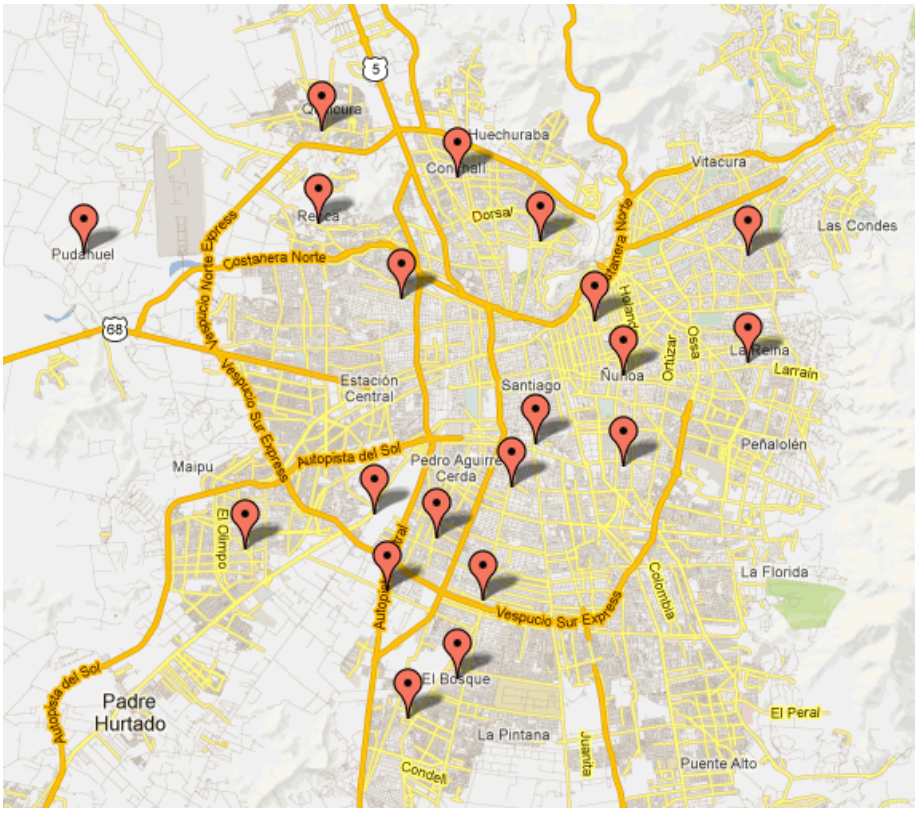
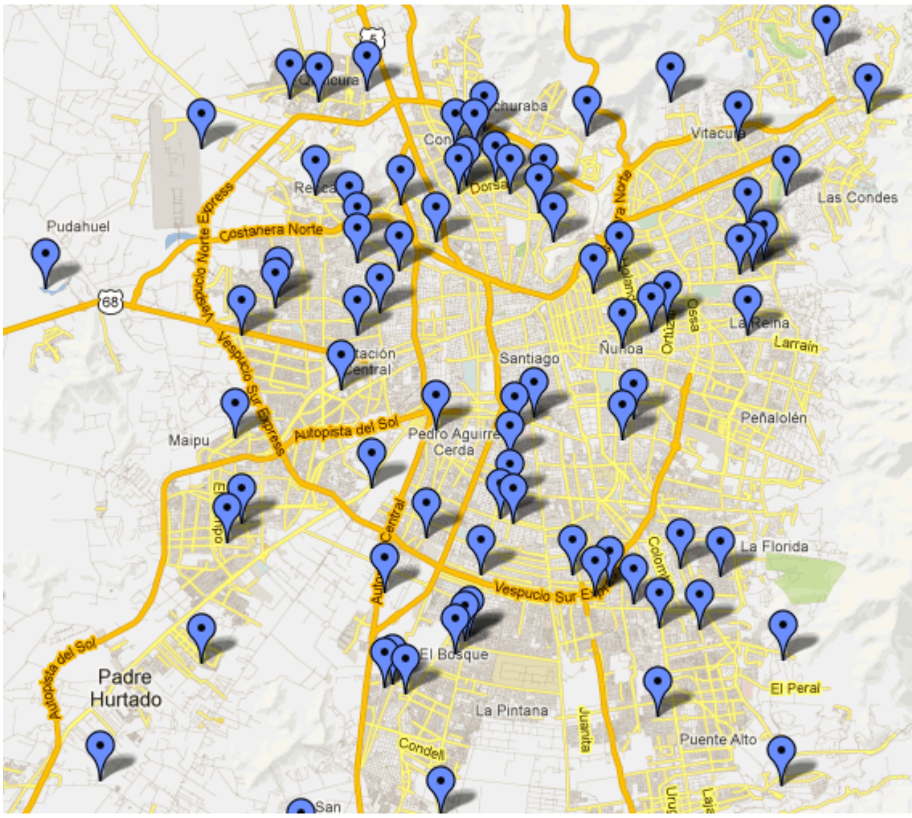


Figure 6: Map of Chilean IP Infrastructure Locations – Santiago City



Map data ©2013 Google

4. Econometric Specification

I estimate two models, a pooled OLS and an instrumental variables regression. I regress the exports of firm i, at location c and time t (X_{it}) on Internet access ($Internet_{ict}$), other controls (Y_{it}) and country-time fixed effects (FE_{kt}) for country k. The Internet access variable represents separately whether the firm has a website or uses email.

$$X_{it} = \beta_0 + \beta_1 Internet_{ict} + \beta_2 Y_{it} + FE_{kt} + \varepsilon_{it} \qquad [1.1]$$

In addition to considering total firm exports, I also consider individually by export mode; direct exports and indirect exports. It is common practice to define exports in terms of log values. Instead, I measure exports as a percentage of sales revenue for two reasons.

Firstly, taking logarithms drops all observations with zero trade flows and therefore the analysis would be conditional on positive trade. Unfortunately, even with a randomly assigned treatment, such regressions no longer have a causal interpretation (Angrist and Pischke, 2008). The measured effect includes the causal effect for adopters that have positive trade and a selection bias. The selection bias represents the change in the composition of the group with positive trade. If the Internet induces marginal firms to start exporting, these firms will have smaller exports than the existing pool of exporters and hence selection bias will be negative.

Secondly, exports are measured as a percentage of sales within the World Bank surveys. The value of exports can be calculated by multiplying this measure by sales. However, the sample size would be reduced due to missing sales data.

Inclusion of country-time fixed effects controls for any country-level factors that might influence firm exports. These would include variables such as WTO membership, GDP, distance from other countries and country-level measures of telecommunications regulations.

I include log values of firm age, the lagged number of employees and lagged labour productivity in US dollars as control variables. Inclusion of additional lagged variables is not possible due to the limited panel aspect of the data. Firm-level research has shown that firms with more skilled workers or that are foreign-owned are more likely to export (Wagner, 2007). However, inclusion of these additional contemporaneous controls is troublesome. It is likely that Internet adoption could lead to firms employing more skilled workers, such as those with IT training, or becoming more attractive to foreign owners.

Therefore these additional controls are unlikely to be exogenous and would introduce selection bias.

In the first stage, I instrument Internet access with IP infrastructure in their locality (IP_{ct}). As a robustness check I introduce additional instruments to capture the effect of IP infrastructure for the very largest firms relative to all other firms. These are defined as firms above the 90th percentile within their country in terms of lagged employees or lagged sales.

$$Internet_{ict} = \gamma_0 + \gamma_1 IP_{ct} + \gamma_2 IP_{ct} * 90percentile_{it-1} + \gamma_4 Y_{it} + FE_{kt} + v_{it} \quad [1.2]$$

The inclusion of country-time fixed effects also absorbs any country-level factors that impact Internet access. I am therefore using within-country variation in IP infrastructure to identify a firm's Internet adoption. The instrumental variable approach quantifies the causal effect of Internet adoption on firm exports for those firms that employ Internet technology because of local variations in infrastructure.

5. Results

I first present results of the regressions involving total firm exports, before considering separately direct exports and indirect exports.

In Table 4, regressions (1) and (4) present the pooled OLS regressions for website and email use respectively. Firms with either measure of Internet access export significantly more of their sales than those that do not. Firms with a website and firms that use email export 4 per cent and 6 per cent more of their sales respectively. Turning to the control variables, as expected larger firms (measured by employees) export significantly more than smaller firms. Oddly, exports appear to be unaffected by productivity. However, more productive firms also tend to be larger and so it is likely this effect is being picked up by the measure of firm size. Younger firms export more than older firms.

Regressions (2)-(3) and (5)-(6) take account of the endogeneity of Internet adoption and present results of the instrumental variable regressions. In order to be a valid instrument, the growth of IP infrastructure needs to predict Internet adoption (relevance) and for the only effect on trade to be through the Internet channel (exogeneity). I examine these conditions in turn.

The first stage regression (2) shows that firms in locations with better IP Infrastructure are more likely to have a website. Specifically, a doubling of IP addresses per person leads to a 13 per cent higher probability of firms in that area adopting a website. The effect is significant at the 1 per cent level. The F-statistic on the instrumental variable is 30.5, which exceeds the commonly-used Staiger and Stock (1997) "rule of thumb" of 10. Again, this suggests IP infrastructure has good explanatory power over the endogenous variable. With instrumental variables, standard errors are biased downwards in small samples. The Stock and Yogo (2005) test uses this bias to detect weak instruments. A bias of less than 10 per cent suggests that the instruments are strong, which is the case here. As one would expect, the control variables indicate that larger and more productive firms are more likely to have a website. The first stage regression (5) shows IP infrastructure is a similarly strong predictor of email use.

It is not possible to test an instrument's exogeneity directly. With more than one instrument it is possible to perform Hansen's (1982) over-identification test. Under the assumption that at least one instrument is valid, the null hypothesis is that the remaining instruments are valid. To test this I introduce the additional (aforementioned) instrument in regressions (3) and (6).

The instrument interacting IP infrastructure with a dummy for the very largest firms (by employment) is highly significant in the first stage. The second instrument has a negative sign, indicating that IP infrastructure is a poorer predictor of Internet access for the very largest firms. Introducing the additional instruments lowers the F-statistic a little, but it remains well above 10. The Stock-Yogo bias is within 10 per cent for email use and 15 per cent for having a website. Crucially, the Hansen test does not reject the null hypothesis of jointly valid instruments in either case. IP infrastructure is a strong predictor of both having a website and using email, with the over-identification tests not able to reject the null of jointly valid instruments.

The second stage regressions (2)-(3) and (5)-(6) suggest that the OLS coefficients are downward biased. Firms that adopt a website because of their local IP infrastructure export 25 per cent more of their sales than those that do not. Similar results are found for email use. Given the extent of missing data and the focus on developing countries, one would expect some measurement error and downward bias. As a comparison, for UK firms Abramovsky and Griffith (2006) find the OLS coefficient for firm Internet usage is downward biased by a factor of between six and twenty.

I next consider the effect of Internet usage on direct exports, noted in Table 5. The pattern is remarkably similar to that of total firm exports. In terms of OLS estimation, firms that have a website (use email) export 4 per cent (5 per cent) more of their sales than those that do not. The instrumental variables provide a strong predictor of website and email adoption, with an F-statistic above 20 in each regression and a Stock-Yogo bias of less than 10 per cent. The Hansen over-identification test cannot reject the null of jointly valid instruments. The second stage regressions suggest that the OLS estimates are again downward biased and that having a website or using email has a positive and highly significant impact on direct exports.

Table 6 reports the impact of Internet usage on indirect exports. The picture here is quite different to that of direct exports or total firm exports. Ignoring the endogeneity of Internet usage, OLS coefficients suggest that having a website has no significant impact on indirect exports. The effect of email, however, is positive and marginally significant.

The instruments, again, perform well as strong predictors of Internet adoption and the Hansen over-identification test cannot reject the null of jointly valid instruments. The second stage regressions suggest that the Internet has a negative yet insignificant effect on indirect exports. This is true of both email and websites.

Table 4: Total Trade Regressions

Dependent Variable:	**Percentage Total Exports / Sales**					
Second Stage	**OLS**	**IV**		**OLS**	**IV**	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Website	**0.041*****	**0.247****	**0.279*****			
	(0.005)	**(0.107)**	**(0.095)**			
Email				**0.059*****	**0.247****	**0.262*****
				(0.006)	**(0.105)**	**(0.086)**
Firm Age	**-0.025*****	**-0.024*****	**-0.025*****	**-0.025*****	**-0.023*****	**-0.022*****
	(0.004)	**(0.004)**	**(0.004)**	**(0.004)**	**(0.004)**	**(0.004)**
Lagged Employees	**0.043*****	**0.018**	**0.005**	**0.043*****	**0.028*****	**0.015***
	(0.002)	**(0.013)**	**(0.012)**	**(0.002)**	**(0.009)**	**(0.008)**
Lagged Productivity	**0.001**	**-0.007**	**-0.007***	**0.001**	**-0.005**	**-0.006***
	(0.002)	**(0.004)**	**(0.004)**	**(0.001)**	**(0.003)**	**(0.003)**
> 90% Lagged Employees			**0.076*****			**0.086*****
			(0.013)			**(0.014)**
Endogenous Variable: **First Stage**		**Website**			**Email**	
IP Infrastructure		**0.132*****	**0.146*****		**0.131*****	**0.151*****
		(0.024)	**(0.025)**		**(0.017)**	**(0.018)**
IP Infrastructure *			**-0.162*****			**-0.231*****
> 90% Lagged Employees			**(0.051)**			**(0.028)**
Firm Age		**-0.002**	**-0.002**		**-0.011****	**-0.011****
		(0.007)	**(0.007)**		**(0.005)**	**(0.005)**
Lagged Employees		**0.121*****	**0.125*****		**0.083*****	**0.092*****
		(0.003)	**(0.004)**		**(0.003)**	**(0.003)**
Lagged Productivity		**0.036*****	**0.036*****		**0.031*****	**0.031*****
		(0.003)	**(0.003)**		**(0.002)**	**(0.002)**
> 90% Lagged Employees			**0.002**			**-0.023**
			(0.021)			**(0.014)**
Weak Identification Test (F statistic)		**30.534**	**19.707**		**48.835**	**39.100**
Stock-Yogo Maximal Bias		**10%**	**15%**		**10%**	**10%**
Hansen Test (p-value)			**0.548**			**0.780**
Country-Year Fixed Effects	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
Observations	**10,812**	**10,812**	**10,812**	**10,827**	**10,827**	**10,827**

(Robust standard errors in parentheses)

Table 5: Direct Trade Regressions

Dependent Variable:	Percentage Direct Exports / Sales					
Second Stage	**OLS**	**IV**		**OLS**	**IV**	
	(7)	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
Website	**0.037***** **(0.005)**	**0.248***** **(0.092)**	**0.295***** **(0.083)**			
Email				**0.054***** **(0.005)**	**0.259***** **(0.093)**	**0.294***** **(0.078)**
Firm Age	**-0.029***** **(0.004)**	**-0.028***** **(0.004)**	**-0.028***** **(0.004)**	**-0.029***** **(0.004)**	**-0.026***** **(0.004)**	**-0.026***** **(0.004)**
Lagged Employees	**0.045***** **(0.002)**	**0.020*** **(0.011)**	**0.006** **(0.011)**	**0.045***** **(0.002)**	**0.028***** **(0.008)**	**0.015**** **(0.008)**
Lagged Productivity	**-0.001** **(0.001)**	**-0.010***** **(0.004)**	**-0.011***** **(0.004)**	**-0.002** **(0.001)**	**-0.008** **(0.003)**	**-0.009***** **(0.003)**
> 90% Lagged Employees			**0.059***** **(0.013)**			**0.074***** **(0.013)**
Endogenous Variable: **First Stage**		**Website**			**Email**	
IP Infrastructure		**0.140***** **(0.024)**	**0.155***** **(0.025)**		**0.134***** **(0.017)**	**0.151***** **(0.018)**
IP Infrastructure * > 90% Lagged Employees			**-0.164***** **(0.050)**			**-0.231***** **(0.028)**
Firm Age		**-0.005** **(0.006)**	**-0.005** **(0.004)**		**-0.015***** **(0.005)**	**-0.011**** **(0.005)**
Lagged Employees		**0.120***** **(0.003)**	**0.123***** **(0.004)**		**0.084***** **(0.002)**	**0.092***** **(0.003)**
Lagged Productivity		**0.038***** **(0.003)**	**0.038***** **(0.003)**		**0.031***** **(0.002)**	**0.031***** **(0.002)**
> 90% Lagged Employees			**0.011** **(0.020)**			**-0.023*** **(0.014)**
Weak Identification Test (F statistic)		**34.738**	**21.984**		**50.981**	**40.332**
Stock-Yogo Maximal Bias		**10%**	**10%**		**10%**	**10%**
Hansen Test (p-value)			**0.308**			**0.504**
Country-Year Fixed Effects	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
Observations	**11,445**	**11,445**	**11,445**	**11,469**	**11,469**	**11,469**

(Robust standard errors in parentheses)

Table 6: Indirect Trade Regressions

Dependent Variable:	Percentage Indirect Exports / Sales					
Second Stage	**OLS**	**IV**		**OLS**	**IV**	
	(13)	**(14)**	**(15)**	**(16)**	**(17)**	**(18)**
Website	**-0.003**	**-0.074**	**-0.066**			
	(0.003)	**(0.046)**	**(0.042)**			
Email				**0.006***	**-0.077**	**-0.062**
				(0.004)	**(0.048)**	**(0.040)**
Firm Age	**-0.006*****	**-0.006*****	**-0.006*****	**-0.006*****	**-0.007*****	**-0.007*****
	(0.002)	**(0.001)**	**(0.002)**	**(0.002)**	**(0.002)**	**(0.002)**
Lagged Employees	**0.005*****	**0.013****	**0.012****	**0.004*****	**0.011*****	**0.009****
	(0.001)	**(0.006)**	**(0.005)**	**(0.001)**	**(0.004)**	**(0.004)**
Lagged Productivity	**-0.002****	**0.001**	**0.004**	**-0.002*****	**0.001**	**0.001**
	(0.001)	**(0.002)**	**(0.005)**	**(0.001)**	**(0.002)**	**(0.001)**
> 90% Lagged Employees			**0.004**			**0.001**
			(0.006)			**(0.006)**
Endogenous Variable:		**Website**			**Email**	
First Stage						
IP Infrastructure		**0.140*****	**0.155*****		**0.134*****	**0.151*****
		(0.024)	**(0.025)**		**(0.017)**	**(0.018)**
IP Infrastructure *			**-0.164*****			**-0.231*****
> 90% Lagged Employees			**(0.050)**			**(0.027)**
Firm Age		**-0.005**	**-0.005**		**-0.015*****	**-0.014****
		(0.006)	**(0.004)**		**(0.005)**	**(0.005)**
Lagged Employees		**0.120*****	**0.123*****		**0.084*****	**0.093*****
		(0.003)	**(0.004)**		**(0.002)**	**(0.003)**
Lagged Productivity		**0.038*****	**0.038*****		**0.031*****	**0.031*****
		(0.003)	**(0.003)**		**(0.002)**	**(0.002)**
> 90% Lagged Employees			**0.011**			**-0.026****
			(0.020)			**(0.013)**
Weak Identification Test (F statistic)		**33.738**	**21.985**		**50.978**	**40.328**
Stock-Yogo Maximal Bias		**10%**	**10%**		**10%**	**10%**
Hansen Test (p-value)			**0.702**			**0.548**
Country-Year Fixed Effects	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
Observations	**11,444**	**11,444**	**11,444**	**11,468**	**11,468**	**11,468**

(Robust standard errors in parentheses)

6. Robustness

In this section I consider the robustness of the instrumental variable results.

Firstly, I restrict the sample size to countries with at least 20 different firm locations. The number of countries in the sample is halved from 20 to 10 (see Table 3) and the number of observations reduced by just over a third. For brevity, I report only results involving both instruments within Table 7.

The instruments continue to perform well. The instruments are strong predictors of Internet usage in the first stage, however, they are slightly weaker compared with the previous section. This is to be expected with smaller sample sizes. Nevertheless, F-statistics remain above 14 and the Stock-Yogo Bias is within 10 per cent and 15 per cent for the email and website regressions respectively. Importantly, the Hansen over-identification test cannot reject the null of jointly valid instruments in each case.

In the second stage, the conclusions with respect to total firm exports and direct exports are largely unchanged. Firms with either measure of Internet access have significantly higher exports relative to those that do not. Conversely, the Internet has a negative and now marginally significant effect on indirect exports.

Secondly, I introduce an alternative second instrument and use the full sample of countries. I employ firm sales instead of employees as a measure of the very largest firms and interact this dummy variable with IP infrastructure. The results are summarised within Table 8.

The alternative second instrument is weaker than the employee specification. For instance, the second instrument is only marginally significant in regression (25). The Stock-Yogo maximal bias is within 10 per cent and 15 per cent for the email and website regressions respectively. The Hansen over-identification test cannot reject the null in all but one of the regressions. In regression (28) the over-identification test rejects the null at the 10 per cent level. In isolation, this would cast some doubt over the validity of the instrument. However, the non-rejection in all other specifications suggests otherwise. The conclusions from the second stage regressions mirror the baseline regressions.

Finally, I consider an alternative definition of IP infrastructure. Previously I used the average number of IP addresses over the year. I now utilise the number of IP addresses at the mid-year instead. The results are presented within Table 9.

The results change very little from before. The instrument remains a strong predictor of Internet use, with a Stock-Yogo bias below 10 per cent in each case. In regression (33) the Hansen over-identification test marginally rejects the null of valid instruments. The null cannot be rejected in all other cases.

The second stage results echo earlier findings. However, both measures of Internet usage have a negative and marginally significant effect on indirect exports.

Table 7: Robustness – Restricted Countries

Dependent Variable:	**Percentage Total Exports / Sales**		**Percentage Direct Exports / Sales**		**Percentage Indirect Exports / Sales**	
Second Stage	**(19)**	**(20)**	**(21)**	**(22)**	**(23)**	**(24)**
Website	**0.209****		**0.307*****		**-0.098***	
	(0.098)		**(0.098)**		**(0.052)**	
Email		**0.224****		**0.329*****		**-0.103***
		(0.105)		**(0.096)**		**(0.055)**
Firm Age	**-0.030*****	**-0.025*****	**-0.023*****	**-0.016*****	**-0.007*****	**-0.009*****
	(0.005)	**(0.005)**	**(0.005)**	**(0.004)**	**(0.002)**	**(0.002)**
Lagged Employees	**0.009**	**0.016**	**-0.008**	**0.003**	**0.017****	**0.014*****
	(0.013)	**(0.010)**	**(0.013)**	**(0.009)**	**(0.007)**	**(0.005)**
Lagged Productivity	**-0.004**	**-0.003**	**-0.007***	**-0.005***	**0.003**	**0.002**
	(0.004)	**(0.003)**	**(0.004)**	**(0.003)**	**(0.002)**	**(0.002)**
> 90% Lagged Employees	**0.088*****	**0.098*****	**0.092*****	**0.106*****	**-0.003**	**-0.007**
	(0.016)	**(0.018)**	**(0.023)**	**(0.016)**	**(0.002)**	**(0.008)**
Endogenous Variable: First Stage	**Website**	**Email**	**Website**	**Email**	**Website**	**Email**
IP Infrastructure	**0.135*****	**0.123*****	**0.135*****	**0.123*****	**0.136*****	**0.123*****
	(0.027)	**(0.019)**	**(0.027)**	**(0.019)**	**(0.026)**	**(0.019)**
IP Infrastructure *	**-0.187*****	**-0.203*****	**-0.187****	**-0.203*****	**-0.187*****	**-0.203*****
> 90% Lagged Employees	**(0.057)**	**(0.028)**	**(0.057)**	**(0.028)**	**(0.057)**	**(0.028)**
Firm Age	**0.004**	**-0.016*****	**0.004**	**-0.016*****	**0.004**	**-0.016*****
	(0.008)	**(0.006)**	**(0.008)**	**(0.006)**	**(0.008)**	**(0.006)**
Lagged Employees	**0.134*****	**0.093*****	**0.134*****	**0.093*****	**0.134*****	**0.093*****
	(0.005)	**(0.004)**	**(0.005)**	**(0.004)**	**(0.005)**	**(0.004)**
Lagged Productivity	**0.033*****	**0.027*****	**0.033*****	**0.027*****	**0.033*****	**0.027*****
	(0.003)	**(0.002)**	**(0.003)**	**(0.002)**	**(0.003)**	**(0.002)**
> 90% Lagged Employees	**-0.007**	**-0.041*****	**-0.007**	**-0.041*****	**-0.007**	**-0.041*****
	(0.027)	**(0.015)**	**(0.027)**	**(0.015)**	**(0.027)**	**(0.015)**
Weak Identification Test (F statistic)	**14.767**	**23.797**	**14.767**	**23.797**	**14,767**	**23.797**
Stock-Yogo Maximal Bias	**15%**	**10%**	**15%**	**10%**	**15%**	**10%**
Hansen Test (p-value)	**0.986**	**0.847**	**0.781**	**0.575**	**0.524**	**0.409**
Country-Year Fixed Effects	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
Observations	**7,198**	**7,210**	**7,198**	**7,210**	**7,198**	**7,210**

(Robust standard errors in parentheses)

Table 8: Robustness – Alternative Second Instrument

Dependent Variable:	**Percentage Total Exports / Sales**		**Percentage Direct Exports / Sales**		**Percentage Indirect Exports / Sales**	
Second Stage	**(25)**	**(26)**	**(27)**	**(28)**	**(29)**	**(30)**
Website	**0.223****		**0.213****		**-0.074**	
	(0.101)		**(0.084)**		**(0.046)**	
Email		**0.180****		**0.178****		**-0.069**
		(0.088)		**(0.078)**		**(0.042)**
Firm Age	**-0.025*****	**-0.023*****	**-0.029*****	**-0.028*****	**-0.006*****	**-0.007*****
	(0.004)	**(0.004)**	**(0.004)**	**(0.004)**	**(0.002)**	**(0.002)**
Lagged Employees	**0.020**	**0.031*****	**0.024****	**0.034*****	**0.014****	**0.011*****
	(0.013)	**(0.008)**	**(0.010)**	**(0.007)**	**(0.006)**	**(0.004)**
Lagged Productivity	**-0.007***	**-0.005**	**-0.008****	**-0.006****	**0.001**	**0.001**
	(0.004)	**(0.003)**	**(0.003)**	**(0.003)**	**(0.002)**	**(0.002)**
> 90% Lagged Sales	**0.011**	**0.023***	**-0.006**	**0.008**	**-0.005**	**-0.011***
	(0.013)	**(0.014)**	**(0.011)**	**(0.013)**	**(0.005)**	**(0.006)**
Endogenous Variable: First Stage	**Website**	**Email**	**Website**	**Email**	**Website**	**Email**
IP Infrastructure	**0.139*****	**0.148*****	**0.148*****	**0.150*****	**0.148*****	**0.151*****
	(0.025)	**(0.018)**	**(0.025)**	**(0.018)**	**(0.025)**	**(0.018)**
IP Infrastructure * > 90% Lagged Sales	**-0.087***	**-0.205*****	**-0.094****	**-0.203*****	**-0.094****	**-0.203*****
	(0.048)	**(0.028)**	**(0.048)**	**(0.028)**	**(0.048)**	**(0.028)**
Firm Age	**-0.002**	**-0.010****	**-0.005**	**-0.013*****	**-0.005**	**-0.014*****
	(0.007)	**(0.005)**	**(0.006)**	**(0.005)**	**(0.006)**	**(0.005)**
Lagged Employees	**0.122*****	**0.090*****	**0.120*****	**0.091*****	**0.120*****	**0.091*****
	(0.003)	**(0.003)**	**(0.003)**	**(0.003)**	**(0.003)**	**(0.003)**
Lagged Productivity	**0.036*****	**0.035*****	**0.038*****	**0.036*****	**0.038*****	**0.036*****
	(0.003)	**(0.002)**	**(0.003)**	**(0.002)**	**(0.003)**	**(0.002)**
> 90% Lagged Sales	**0.013**	**-0.033****	**0.020**	**-0.038*****	**0.020**	**-0.038*****
	(0.021)	**(0.014)**	**(0.021)**	**(0.013)**	**(0.021)**	**(0.013)**
Weak Identification Test (F statistic)	**16.456**	**35.430**	**18.743**	**36.535**	**18.743**	**36.533**
Stock-Yogo Maximal Bias	**15%**	**10%**	**15%**	**10%**	**15%**	**10%**
Hansen Test (p-value)	**0.456**	**0.203**	**0.210**	**0.071**	**0.892**	**0.732**
Country-Year Fixed Effects	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
Observations	**10,812**	**10,827**	**11,445**	**11,469**	**11,444**	**11,468**

(Robust standard errors in parentheses)

Table 9: Robustness - Mid-Year IP Infrastructure

Dependent Variable:	**Percentage Total Exports / Sales**		**Percentage Direct Exports / Sales**		**Percentage Indirect Exports / Sales**	
Second Stage	**(31)**	**(32)**	**(33)**	**(34)**	**(35)**	**(36)**
Website	**0.236***** **(0.091)**		**0.261***** **(0.079)**		**-0.073*** **(0.042)**	
Email		**0.234***** **(0.084)**		**0.274***** **(0.076)**		**-0.068*** **(0.041)**
Firm Age	**-0.025***** **(0.004)**	**-0.022***** **(0.004)**	**-0.028***** **(0.004)**	**-0.026***** **(0.004)**	**-0.006***** **(0.002)**	**-0.007***** **(0.002)**
Lagged Employees	**0.010** **(0.012)**	**0.018**** **(0.008)**	**0.010** **(0.010)**	**0.017**** **(0.007)**	**0.012**** **(0.005)**	**0.010**** **(0.004)**
Lagged Productivity	**-0.006** **(0.004)**	**-0.005** **(0.003)**	**-0.010***** **(0.003)**	**-0.008***** **(0.003)**	**0.001** **(0.002)**	**0.001** **(0.002)**
> 90% Lagged Employees	**0.074***** **(0.014)**	**0.084***** **(0.014)**	**0.058***** **(0.013)**	**0.073***** **(0.013)**	**0.004** **(0.006)**	**0.001** **(0.006)**
Endogenous Variable: First Stage	**Website**	**Email**	**Website**	**Email**	**Website**	**Email**
IP Infrastructure	**0.159***** **(0.026)**	**0.161***** **(0.019)**	**0.168***** **(0.026)**	**0.164***** **(0.019)**	**0.168***** **(0.050)**	**0.164***** **(0.019)**
IP Infrastructure * > 90% Lagged Employees	**-0.147***** **(0.050)**	**-0.242***** **(0.029)**	**-0.150***** **(0.050)**	**-0.241***** **(0.028)**	**-0.150**** **(0.050)**	**-0.241***** **(0.028)**
Firm Age	**-0.002** **(0.007)**	**-0.011**** **(0.005)**	**-0.005** **(0.006)**	**-0.014***** **(0.005)**	**-0.004** **(0.006)**	**-0.014***** **(0.005)**
Lagged Employees	**0.125***** **(0.004)**	**0.092***** **(0.003)**	**0.123***** **(0.004)**	**0.093***** **(0.003)**	**0.123***** **(0.004)**	**0.093***** **(0.003)**
Lagged Productivity	**0.036***** **(0.003)**	**0.031***** **(0.002)**	**0.038***** **(0.003)**	**0.031***** **(0.002)**	**0.038***** **(0.003)**	**0.031***** **(0.002)**
> 90% Lagged Employees	**-0.002** **(0.021)**	**-0.022** **(0.014)**	**0.008** **(0.020)**	**-0.025*** **(0.013)**	**0.008** **(0.020)**	**-0.025*** **(0.013)**
Weak Identification Test (F statistic)	**20.439**	**40.687**	**22.723**	**41.930**	**22.723**	**41.927**
Stock-Yogo Maximal Bias	**10%**	**10%**	**10%**	**10%**	**10%**	**10%**
Hansen Test (p-value)	**0.238**	**0.451**	**0.096**	**0.348**	**0.835**	**0.601**
Country-Year Fixed Effects	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
Observations	**10,812**	**10,827**	**11,445**	**11,469**	**11,444**	**11,468**

(Robust standard errors in parentheses)

7. Conclusions

In this paper I examine the effects of Internet adoption on firm-level exports in developing countries. In particular, I consider how the Internet affects a firm's choice of export channel, namely, whether a firm exports directly or via intermediaries. A significant contribution is the construction of a novel instrumental variable for firm Internet adoption. I use within-country geographic variation in IP addresses to identify the causal effect of Internet use on exports. I combine the instrument with firm location information to provide an exogenous predictor of Internet use.

Using cross-country analysis, I find that firms with a website, or those that use email, have higher exports than those that do not. Firms with Internet access have greater direct exports, whereas there is a limited and, if anything, negative effect for indirect exports. Importantly, these results are robust to consideration of endogeneity.

The role of intermediaries is particularly prominent within developing countries. With Internet infrastructure rapidly spreading across the developing world, these results suggest the Internet may provide an alternative mechanism to access foreign markets.

8. References

Abel-Koch, J., (2011). Who uses intermediaries in international trade? Evidence from firm-level survey data. *GEP Research Paper Series* 2011/25.

Abramovsky, L., Griffith, R., (2006). Outsourcing and offshoring of business services: How important is ICT? *Journal of European Economic Association* 4(2-3), pp594-601.

Ahn, J., Khandelwal, A. K., S-J. Wei (2011). The role of Intermediaries in facilitating trade. *Journal of International Economics* 84(1), pp73-85.

BBC News, (2012). No borders mean flatter world will get us anywhere.

[Accessed 10 February 2013]. Available from: http://www.bbc.co.uk/news/business-18594623

Bernard, A. B., Eaton, J., Jensen, J. B., P. K. Schott (2007). Firms in International Trade. *Journal of Economic Perspectives* 21(3), pp 105-130.

Bernard, A. B., Jensen, J. B., P. Schott (2009). Importers, exporters and multinationals: a portrait of firms in the U.S. that trade goods. T. Dunne, J.B. Jensen, M.J. Roberts (eds), Producer Dynamics: New Evidence from Micro Data. Yale University Press.

Bernard, A. B., Jensen, J. B., Redding, S., P. K. Schott (2010). Wholesalers and retailers in US trade. *American Economic Review* 100(2), pp408-413.

Blum, B., Claro, S., I. Horstmann, (2010). Facts and figures on intermediated trade. *American Economic Review* 100(2), pp419-423.

Clarke, G., Wallsten, S., (2006). Has the Internet increased trade? Developed and Developing Country evidence. *Economic Inquiry* 44(3), pp 465-484.

Clarke, G., (2008). Has the Internet increased exports for firms from low and middle-income countries? *Information Economics and Policy* 20, pp 16-37.

EU, (2013). Digital Agenda For Europe. [Accessed 1 February 2013]. Available from: http://ec.europa.eu/digital-agenda/

Freund, C., Weinhold, D., (2002). The Internet and international trade in services. *American Economic Review* 92(2), pp 236-240.

Freund, C., Weinhold, D., (2004). The effect of the Internet on international trade. *Journal of International Economics*, 62(1), pp 171-189.

Government of Kenya, (2013). Vision 2030. [Accessed 1 February 2013]. Available from: http://www.vision2030.go.ke/

Haller, S., Siedschlag, I., (2011). Determinants of ICT adoption: evidence from firm-level data. *Applied Economics* 43, pp 3775-3788.

Hansen, L., (1982). Large sample properties of generalised method of moments estimators. *Econometrica* 50, pp 1029-1054.

Maxmind, (2013). GeoIP City Accuracy for Selected Countries. [Accessed 01 February 2013]. Available from http://www.maxmind.com/en/city_accuracy

Ricci, L. A., Trionfetti, F., (2012). Productivity, Networks, and Export Performance: Evidence from a Cross-country Firm Dataset.

Staiger, D., Stock, J., (1997). Instrumental variables regression with weak instruments. *Econometrica* 65, pp 557-586.

Stock, J., Yogo, M., (2005). Testing for weak instruments in linear IV regression, identification and inference for econometric models. Essays in honour of Thomas Rothenberg. Ed. Andrews, D., Stock, J., pp 80-108.

Tang, L., (2006). Communication costs and Trade of Differentiated Goods. *Review of International Economics,* 14(1), pp 54-68.

UNESCO, (2013). Broadband Commission. [Accessed 1 February 2013]. Available from: http://www.broadbandcommission.org/

Wagner, J. (2007). Exports and productivity: A survey of the evidence from firm-level Data. *The World Economy* 30(1), pp 60-82.