

ICT and Education: Evidence from Pupil Home Addresses

Rosa Sanchis-Guarner (LSE-SERC-Grantham), with Ben Faber (UC Berkeley),
Felix Weinhardt (LSE-CEP, IZA)

8th IZA Conference on Labor Market Policy Evaluation, 18-10-2013

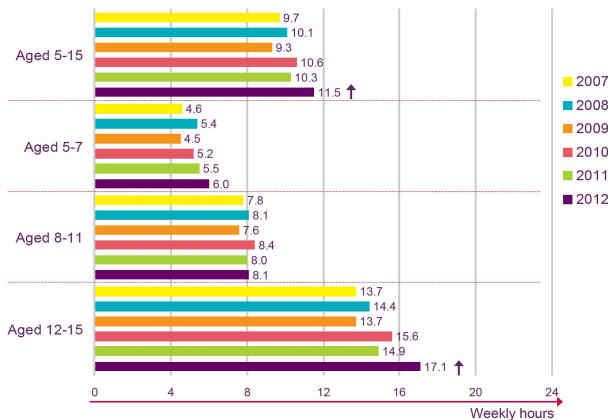
Motivation

- Governments around the world spend a lot of tax money on increases in internet speed.
 - European Budget announced October 2012: spend 7bn Euro to increase broadband speed. UK plans to make superfast broadband (25-30Mbps+) available to 90% of population (and “basic broadband” for all) by 2015. Total cost: £2.4bn.
 - In 2011, the Government set aside £100 million for an Urban Broadband Fund (UBF) to create up to ten ‘super-connected’ cities across the UK.
- However, existing empirical evidence about the socio-economic consequences of falling information and communication costs is very limited.

Broadband and education

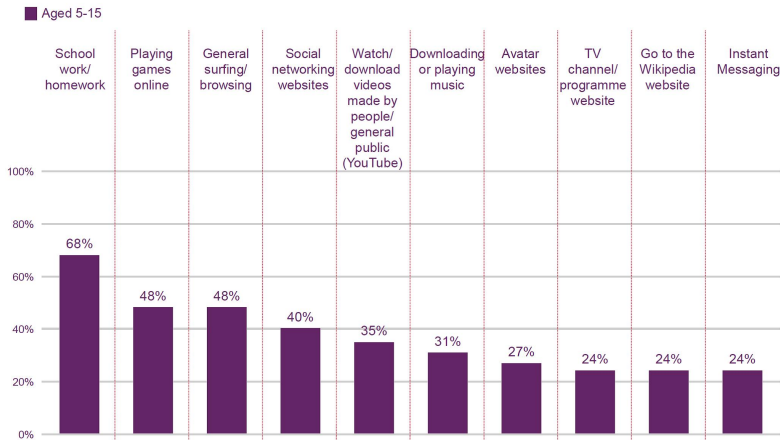
- “[A]bsent this connectivity, individuals cannot take advantage of opportunities for education and skills development [...]” (The White House, Report June 13)

Figure 36: Estimated weekly hours of internet consumption, by age: 2007, 2008, 2009, 2010, 2011 and 2012



QP29A-B/ QC13A-B – How many hours would you say he/ she spends using the internet at home on a typical school day/ on a weekend day? (spontaneous question, single coded)

Children and the Internet, Ofcom Survey 2012



This Paper

- Provide empirical evidence on the following question of policy interest:
 - **How do falling information and communication costs affect educational attainment?**
- Provide empirical evidence from linking broadband speeds with rich administrative data covering the population of English pupil cohorts 2002-2008.

Broadband and Knowledge, the channel

- ICT can affect learning outcomes by changing the productivity (efficiency shifter):
 - e.g. “Wikipedia” increases the efficiency of learning per hour of studies.
- ICT can affect learning outcomes by changing studying time:
 - e.g. “Facebook” or online gaming shift pupil time allocation away from studies.
- ICT can increase in learning effort by making it more attractive relative to other time uses:
 - “Lumosity” effect

Preview of (Preliminary) Findings

- Empirical strategy:
 - Exploit plausibly exogenous variation in internet speed and availability across pupil residences.
- Preview of findings:
 - Higher internet speed has a significant positive effect on age-14 and age-16 national pupil test scores.

Preview of the strategy

Our aim is to estimate the effect of ICT on educational outcomes. But no information on pupils' access or use of internet.

We proceed in three steps:

- **STEP 1:** We estimate the technological relationship between distance to exchange station and postcode average internet speed
- **STEP 2:** We estimate the effect of the distance to exchange station on pupils test scores. We explore heterogeneity
- **STEP 3:** We (attempt to) quantify the effects

Outline of Talk

- Related Literature
- Theoretical framework
- **STEP 1:** Technological relationship between speed and distance
- **STEP 2:** Reduced form identification strategy
 - Main results
 - Placebo
 - Heterogeneity
- **STEP 3:** Quantification of the effects
- Conclusions

The socio-economic effects of internet

- Paper relates to growing empirical literature on the socio-economic consequences of the internet:
 - Effect on voting:
 - (Germany) by: Falck, Gold and Heblich (IZA DP 2012)
 - (Italy) by: Campante, Durante and Sobbrino (NBER WP 2013).
 - Effect on sex crime (Norway) by: Bhuller, Havnes, Leuven and Mogstad (RES *forthcoming*).
 - Effect on social behaviour (Germany) by: Bauernschuster, Falck and Woessmann (IZA DP 2011).
 - Effects on Economic Growth (OECD) by: Czernich, Falck, Kretschmer and Woessmann (EJ 2011).
 - Effect on marriage markets (USA) by: Bellou (IZA DP 2013).
 - Effect on labour markets (USA) by: Forman et al (wages, AER 2012), Atasoy (employment, ILRReview 2013)

ICT and Education

- A few papers have studied the effect of ICT on educational outcomes:
 - Computers in school by: Barrera-Osorio and Linden (WB, 2009)
 - Broadband in school by: Belo et al (Portugal - WP 2011)
 - Human Capital and computers at home by: Fairlie and London (EJ 2012), Malamud et al (QJE 2013), Beuermann et al (NBER 2013)

ICT and educational outcomes

$$H_i = A_i L_i^\beta$$

- where H_i is educational achievement, A_i is a pupil specific productivity shifter, L_i are study hours, and $\beta > 0$ is the elasticity of achievement with respect to time spent studying

$$L_i = e^{\rho_i} w_i^\gamma$$

- where ρ_i is a pupil specific study hours supply shifter, w_i is the pay-off for study hours, and $\gamma > 0$ is the elasticity of study hours with respect to the relative attractiveness of studying

Estimation equation

Information and communication costs c_i

$$\tilde{c}_i = \frac{1}{c_i}$$

$$A_{ins} = \tilde{c}_i^\delta e^{\lambda_i \phi_n \mu_s \varepsilon_{ins}}$$

and

$$w = \tilde{c}_i^\eta$$

hence:

$$\ln H_{ins} = \alpha + \tilde{\beta} \ln \tilde{c}_i + \phi_n + \mu_s + \theta_i + \varepsilon_{ins}$$

where

$$\tilde{\beta} = (\delta + \beta \gamma \eta)$$

and

$$\theta_i = \beta \rho_i + \lambda_i$$

Hypothesis

$$\ln H_{ins} = \alpha + \tilde{\beta} \ln \tilde{c}_i + \phi_n + \mu_s + \theta_i + \varepsilon_{ins}$$

Hypothesis:

“*Wikipedia* effect” $\delta > 0$ ”

“*Facebook* effect” $\eta < 0$ and “*Lumosity* effect” $\eta > 0$ ”

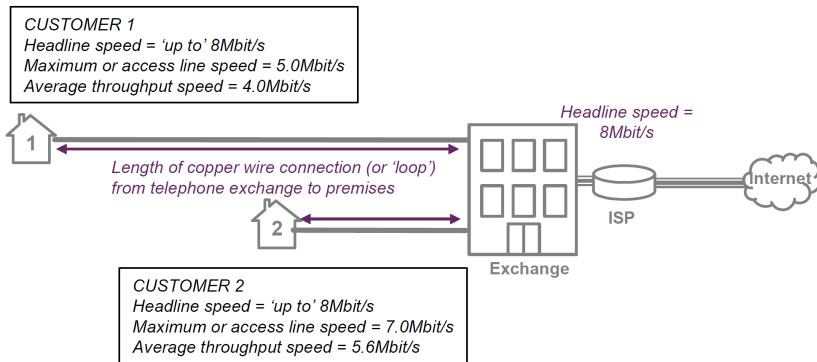
Identification assumption: ε_{ins} and $\ln \tilde{c}_i$ are uncorrelated

From theory to empirics

- We want to estimate the effect of a reduction of ICT costs on educational outcomes, but
 - we do not observe pupils ICT costs
- Speed data is available to us but with limitations
 - Not all English postcodes, available only for 2012 and measured with noise
- There is a technological relationship between speed and distance to telephone exchange station
 - This information is available for all pupils locations
 - We exploit this to assess the validity of a reduced-form approach (**Step 1**)
- We estimate the reduced form approach with emphasis on identification of causal effects (**Step 2**)
- We combine results from steps 1 and 2 and quantify the effects (**Step 3**)

STEP 1: Technological relationship between speed and distance

The Technological Distance Relationship



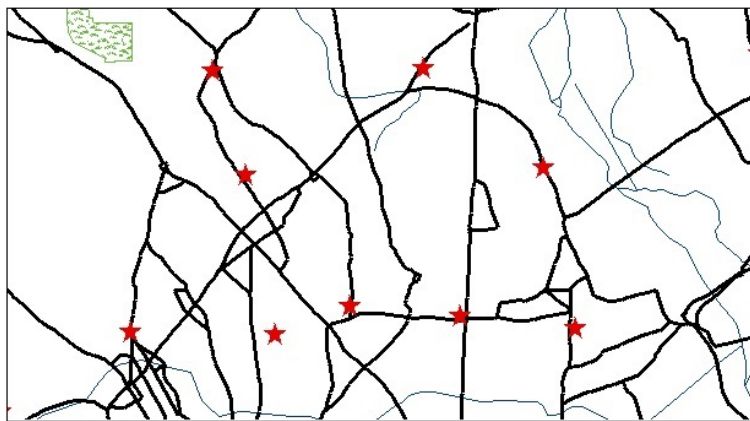
Source: Ofcom Report UK Broadband Speeds 2008

Data Sources

To estimate the distance-speed relationship we use several data sources:

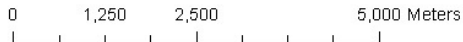
- Location of telephone exchange stations (OLO): PointTopic.
- Postcode level internet speed 2012: OfCom.
 - Available for a sub-set of postcodes (about 55% sample)
 - Median, Average, Maximum speed (30 max)
- Postcode-stations mapping and number of landlines: SamKnows.

Exchange Location



- ★ Exchange Station
- Major Road
- River
- Woodland

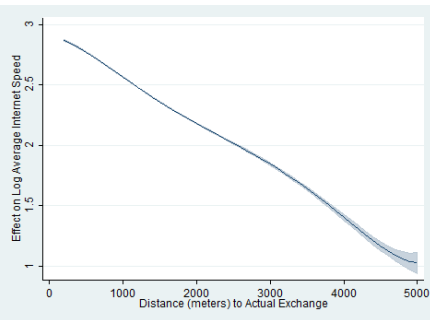
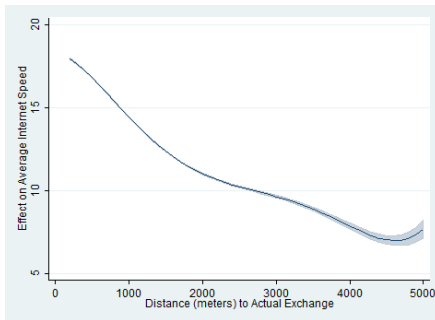
NORTH LONDON



Descriptives First Stage

Average postcode speed, 30 if ≥ 30			
Distance to nearest exchange station (kms)	Mean	Std. Dev.	Freq.
0 to 1km	15.829	5.859	205,469
1 to 2kms	13.092	6.694	223,698
2 to 3kms	10.591	7.691	108,372
3 to 4kms	8.244	8.150	37,443
4 to 5kms	5.896	7.851	9,259
Total	13.166	7.180	584,241

The Technological First Stage Relationship



First Stage Estimation

$$\ln Speed_p = \alpha + \beta DistExchange_p + \varepsilon_p$$

Depvar: Log average postcode internet speed	(1)	(2)	(3)
Distance to connected exchange station (kms)	-0.377*** (0.00266)	-0.404*** (0.00232)	-0.379*** (0.00199)
OLO characteristics		Yes	No
OLO FX			Yes
Geographical controls			
Observations	584,241	584,241	584,241
R-squared	0.243	0.352	0.557

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Clustered at postcode level

STEP 2: Reduced form identification strategy

Specification

$$TestScore_{iprst} = \beta DistExchange_p + \gamma X_{iprt} + \alpha_r + \alpha_s + \alpha_{tKS} + \varepsilon_{iprst}$$

- $DistExchange_p$ = distance to connected exchange station
- X_{iprt} = vector of pupil and postcode controls
- $TestScore_{iprst}$ = Pupil national test score
- p = postcode
- r = exchange station
- s = school
- t = year
- KS = key stage (3 or 4)

Data Sources

- NPD/PLASC: 2002 to 2008 pupil data including residential postcode, gender, ethnicity, FSME.
- Age-7, age-14 and age-16 national Key-Stage test score outcomes. (which we percentalise by cohort, test and subject)
- Postcode-level geographical characteristics: distance to (dis-)amenities, from Ordnance Survey, CBRE.
- Land Registry house price data from mid-2000 until 2002, detrended by 6-month seasons (all housing transactions in England): create postcode-level price-index.
- For the main analysis we group age-14 and age-16 tests in order to have more observations in our (small) postcode-cells.

Descriptives

Variable	Obs	Mean	Std. Dev.	Min	Max
Score	5,824,765	47.452	25.947	1	100
Score_English	5,824,765	46.818	28.119	1	100
Score_Maths	5,824,765	48.146	27.974	1	100
Score_Science	5,824,765	47.392	28.151	1	100
Distance to connected exchange station (kms)	5,824,765	1.520	0.901	0.001	4.999
Distance to nearest exchange station (kms)	5,824,765	1.427	0.809	0.001	4.999
Distance to nearest school (prim/second)	5,824,765	0.461	0.408	0.001	8.594
Distance to nearest road (motorway/A/B)	5,824,765	0.396	0.450	0	14.260
Distance to nearest supermarket	5,824,765	0.869	1.094	0.001	26.943
Distance to nearest station (rail/tube)	5,824,765	1.898	2.709	0	60.902
Distance to nearest tourist office	5,824,765	5.974	4.482	0.002	77.444
Distance to nearest public garden	5,824,765	12.222	7.047	0.030	58.710
Price Deviation (2000-02)	5,824,765	-0.041	0.664	-2.441	4.605
Ethnicity White British	5,824,720	0.823	0.382	0	1
Ethnicity Other white	5,824,720	0.021	0.144	0	1
Ethnicity Asian	5,824,720	0.066	0.247	0	1
Ethnicity Black	5,824,720	0.030	0.171	0	1
Ethnicity Chinese	5,824,720	0.004	0.060	0	1
Ethnicity Mixed	5,824,720	0.020	0.139	0	1
Ethnicity Other	5,824,720	0.009	0.093	0	1
Ethnicity Unknown	5,824,720	0.028	0.166	0	1
FSME Status	5,824,206	0.893	0.309	0	1
Males	5,824,765	0.498	0.500	0	1

OLO/postcode/year cells

Variable	Obs	Mean	Std. Dev.	Min	Max
<i>Per postcode-OLO-year cell</i>					
Number of pupils	2,779,772	2.095	1.460	1	22
Number of pupils KS3	2,779,772	1.077	1.020	0	14
Number of pupils KS4	2,779,772	1.019	0.987	0	13
<i>Per postcode all years</i>					
Number of pupils	639,269	9.112	7.940	1	115
Number of pupils KS3	639,269	4.682	4.213	0	65
Number of pupils KS4	639,269	4.429	3.992	0	50
Number of OLOs	639,269	4.348	1.991	1	7
<i>Per OLO-year</i>					
Number of pupils	27,140	214.619	307.551	1	2798
Number of pupils KS3	27,140	110.285	159.099	0	1469
Number of pupils KS4	27,140	104.335	149.286	0	1337
Number of postcodes	27,140	102.423	138.632	1	1187
<i>Per OLO all years</i>					
Number of pupils	3,912	1488.948	2139.131	1	18713
Number of pupils KS3	3,912	765.114	1104.917	0	9726
Number of pupils KS4	3,912	723.834	1035.893	0	8987
Number of postcodes	3,912	101.658	137.992	1	1163.29

Identification of β

$$TestScore_{iprst} = \beta DistExchange_p + \gamma X_{iprt} + \alpha_r + \alpha_s + \alpha_{tKS} + \varepsilon_{iprst}$$

- Identifying assumption:
 - Distance to exchange does not affect test scores other than through broadband speed, conditional on controls.
- Robustness:
 - Always include exchange-station to control for local neighbourhood conditions and school fixed effects.
 - Observe point estimates before and after including a rich set of controls.
 - Placebo: 15% of postcodes not connected to closest exchange for historical reasons.
 - Test heterogeneity of effect with respect to pupil gender and school meal eligibility (FSME), which allows controlling for distance.

Identification - main worry

- Where would you place exchange stations to maximise coverage?
- placed for voice telephony (copper wire) during 1940s. Distance/length of cable has no impact on quality of connection.
 - likely to be indirectly correlated with geographical features, population density, etc
 - if these (potentially unobserved) factors correlate with house prices ->omitted variable bias through this 'passive sorting'

Score and distance

Depvar: Total KS3/KS4 Score (percentile)	(1)	(2)	(3)	(4)	(5)
Distance to connected exchange station (kms)	0.389*** (0.0248)	0.306*** (0.0214)	-0.182*** (0.0246)	-0.248*** (0.0220)	-0.259*** (0.0215)
Year-stage FX	Yes	Yes	Yes	Yes	Yes
OLO FX	Yes	Yes	Yes	Yes	Yes
School FX		Yes	Yes	Yes	Yes
Geography controls			Yes	Yes	Yes
House price control				Yes	Yes
Pupil controls					Yes
Observations	5,824,765	5,824,765	5,824,765	5,824,765	5,824,161
R-squared	0.100	0.199	0.234	0.263	0.276

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Clustered at postcode level

Score and distance - Placebo

Depvar: Total KS3/KS4 Score (percentile)	(1)	(2)	(3)
Distance to connected exchange station (kms)	-0.259*** (0.0215)		-0.196*** (0.0522)
Distance to nearest exchange station (kms)		-0.274*** (0.0237)	-0.0765 (0.0575)
Year-stage FX	Yes	Yes	Yes
OLO FX	Yes	Yes	Yes
School FX	Yes	Yes	Yes
Geogr, HP and pupil controls	Yes	Yes	Yes
Observations	5,824,161	5,824,161	5,824,161
R-squared	0.276	0.276	0.276

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Clustered at postcode level

Heterogeneity of ICT Effect - Gender

Depvar: Total KS3/KS4 Score (percentile)	(MALE)	(FEMALE)	(INTERACT)
Distance to connected exchange station (kms)	-0.236*** (0.0274)	-0.284*** (0.0276)	
Male*distance			0.0515* (0.0279)
Year-stage FX	Yes	Yes	Yes
AREA FX	OLO	OLO	POSTCODE
School FX and controls	Yes	Yes	Yes
Observations	2,920,806	2,903,355	5,824,161
R-squared	0.276	0.277	0.435
Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Clustered at postcode level			

Heterogeneity of ICT Effect - Income

Depvar: Total KS3/KS4 Score (percentile)	(FSME)	(NON FSME)	(INTERACT)
Distance to connected exchange station (kms)	-0.00592 (0.0582)	-0.285*** (0.0224)	
FMSE*distance			0.0817 (0.0537)
Year-stage FX	Yes	Yes	Yes
AREA FX	OLO	OLO	POSTCODE
School FX and controls	Yes	Yes	Yes
Observations	624,389	5,199,772	5,824,161
R-squared	0.157	0.259	0.435

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Clustered at postcode level

STEP 3: Quantification of the effects

IV Results

Depvar: Total KS3/KS4 Score (percentile)	(OLS)	(RED F)	(IV)
Log average postcode internet speed	0.113** (0.0451)		1.786*** (0.0672)
Distance to connected exchange station (kms)		-0.496*** (0.0298)	
Year-stage FX	Yes	Yes	Yes
OLO FX	Yes	Yes	Yes
All controls	Yes	Yes	Yes
F-stat 1 st stage			445716.5
Observations	3,508,856	3,508,856	3,508,848
R-squared	0.185	0.185	0.126

Quantification attempt

- Doubling speed associated with 1.8 percentile point increase in age-14/age-16 national test scores in this IV.
- Living 4km away from exchange, rather than on exchange: 10% of a standard deviation increase in national test scores.
 - comparable to (lower-bound estimate of) being taught by a teacher one-standard deviation better than the average for a whole year (Aaronson, et al. 2007; Rivkin et al. 2005)
- On our to-do list: Benchmark broadband infrastructure investment against other education interventions where we have reliable estimates, i.e. class-size reductions.
- Note: This IV result is larger than 'reduced form/first stage' from Steps 1 and Step 2 above, which would suggest an effect size of half this size -but based on different samples.

Conclusion

- We invest a great deal in internet speed.
- New evidence of impact of access to fast broadband on education outcomes: doubling speed increases test scores by 7% of standard deviation.
- Combining a number of detailed geographical datasets: geo-coded administrative student data, geo-coded postcode-to-exchange station mapping, universe of English house price transactions, GIS-computed postcode-level amenity variables.
- Careful strategy to deal with endogenous placement/persistence in Geography that could bias estimates exploiting speed-distance relationship.
- Interesting effect heterogeneity by income: contribution to widening gap in income/education relation.

• **Thank You !**

Heterogeneity of ICT Effect - year

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Depvar: Total KS3/KS4 Score (percentile)	Year 2002	Year 2003	Year 2004	Year 2005	Year 2006	Year 2007	Year 2008
Distance to connected exchange station (kms)	-0.366*** (0.0405)	-0.366*** (0.0405)	-0.236*** (0.0394)	-0.323*** (0.0392)	-0.203*** (0.0390)	-0.258*** (0.0396)	-0.251*** (0.0401)
Year-stage FX	Yes	Yes	Yes	Yes	Yes	Yes	Yes
OLO FX	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FX and controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	790,325	790,325	828,209	848,940	862,401	842,135	824,923
R-squared	0.293	0.293	0.291	0.291	0.287	0.282	0.280

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Clustered at postcode level

Heterogeneity of ICT Effect - subject

	(1)	(2)	(3)
Depvar: Total KS3/KS4 Score (percentile)	English	Maths	Science
Distance to connected exchange station (kms)	-0.350*** (0.0225)	-0.207*** (0.0230)	-0.221*** (0.0229)
Year-stage FX	Yes	Yes	Yes
OLO FX	Yes	Yes	Yes
School FX and controls	Yes	Yes	Yes
Observations	5,824,161	5,824,161	5,824,161
R-squared	0.260	0.240	0.251

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Clustered at postcode level

Heterogeneity of ICT Effect - age

	(1)	(2)	(3)
Depvar: KS1 Score (percentile)	Total KS1	Maths KS1	English KS1
Distance to connected exchange station (kms)	-0.413*** (0.0557)	-0.203*** (0.0283)	-0.210*** (0.0307)
Year FX	Yes	Yes	Yes
OLO FX	Yes	Yes	Yes
School FX and controls	Yes	Yes	Yes
Observations	3,095,827	3,095,827	3,095,827
R-squared	0.162	0.129	0.178

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Clustered at postcode level