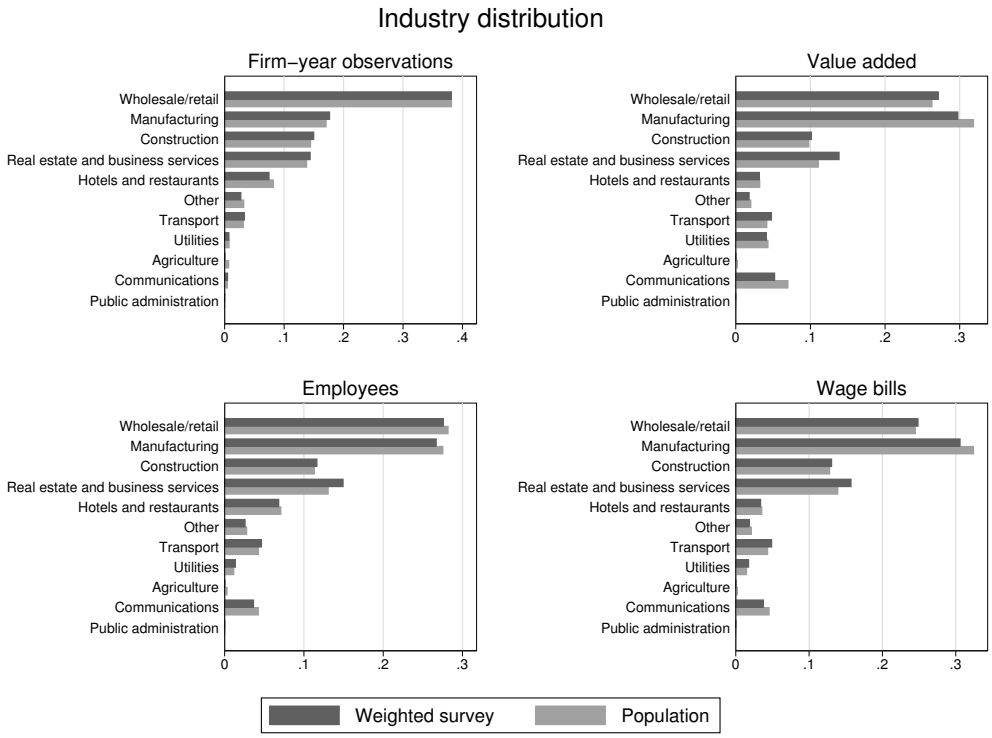
THE SKILL COMPLEMENTARITY OF BROADBAND INTERNET: ONLINE APPENDIX

Anders Akerman, Ingvil Gaarder and Magne Mogstad

May 22, 2015

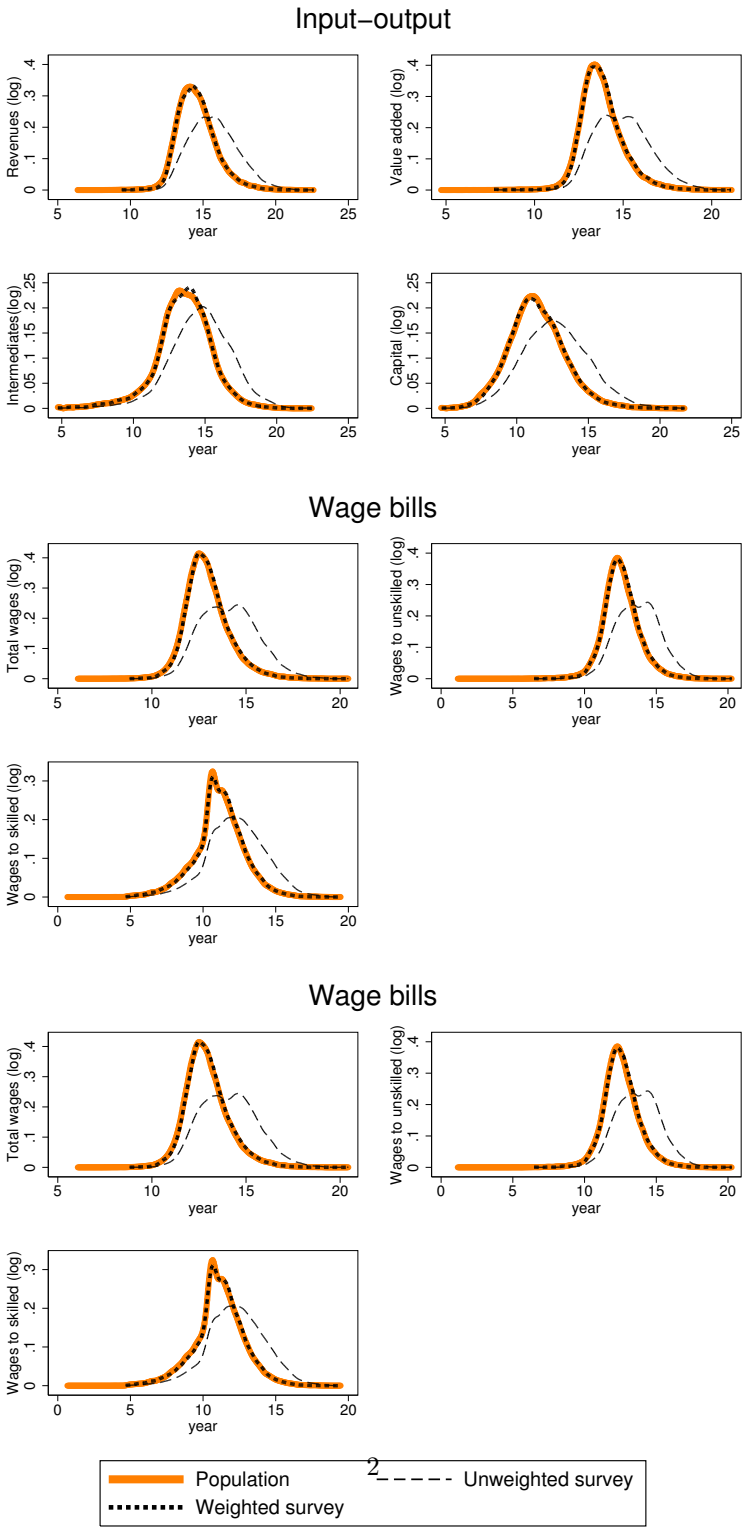
Appendix A: Data and expansion

Figure A.1: Distribution of firms by industry



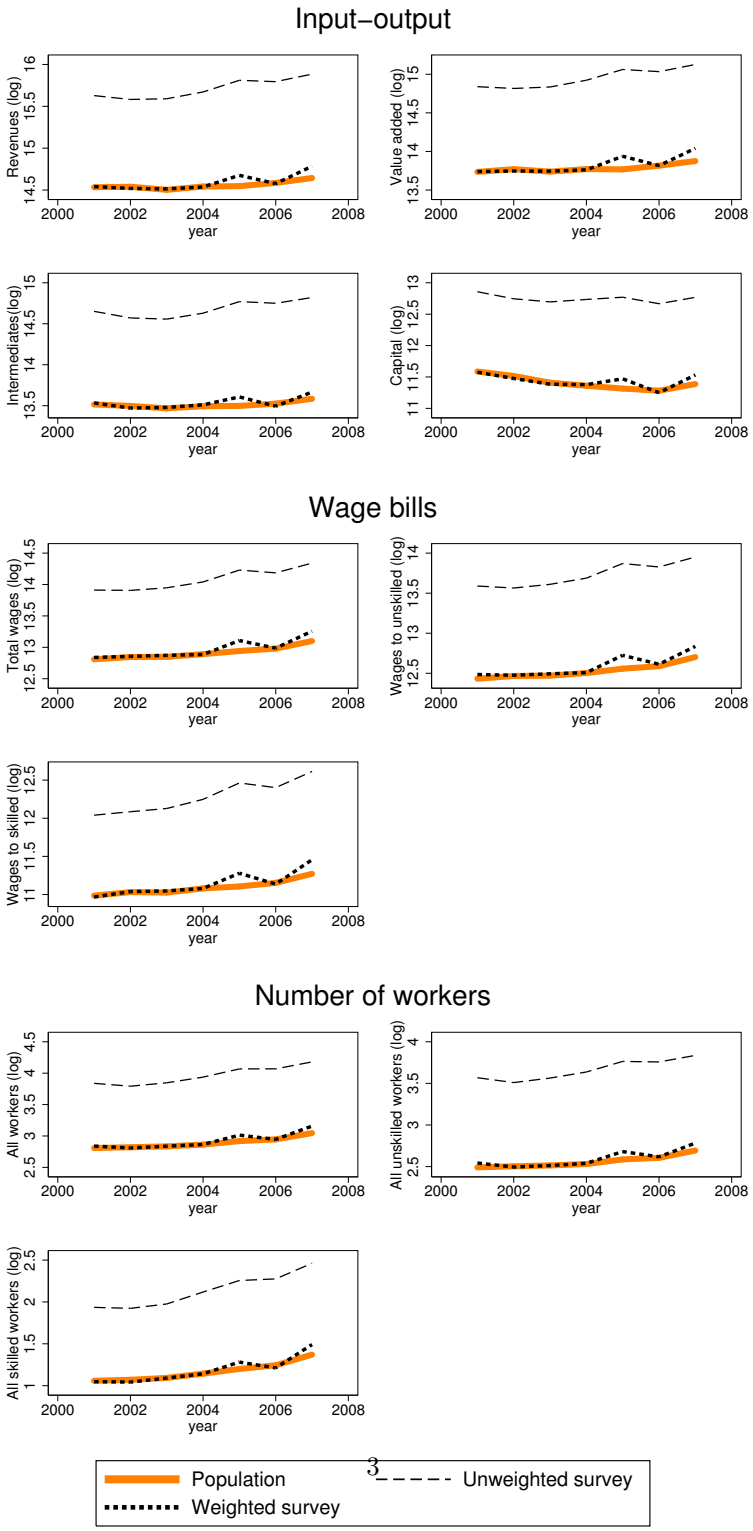
Note: The figure compares the weighted survey sample of joint-stock firms to the population of joint-stock firms.

Figure A.2: Cross-sectional distribution of key firm variables



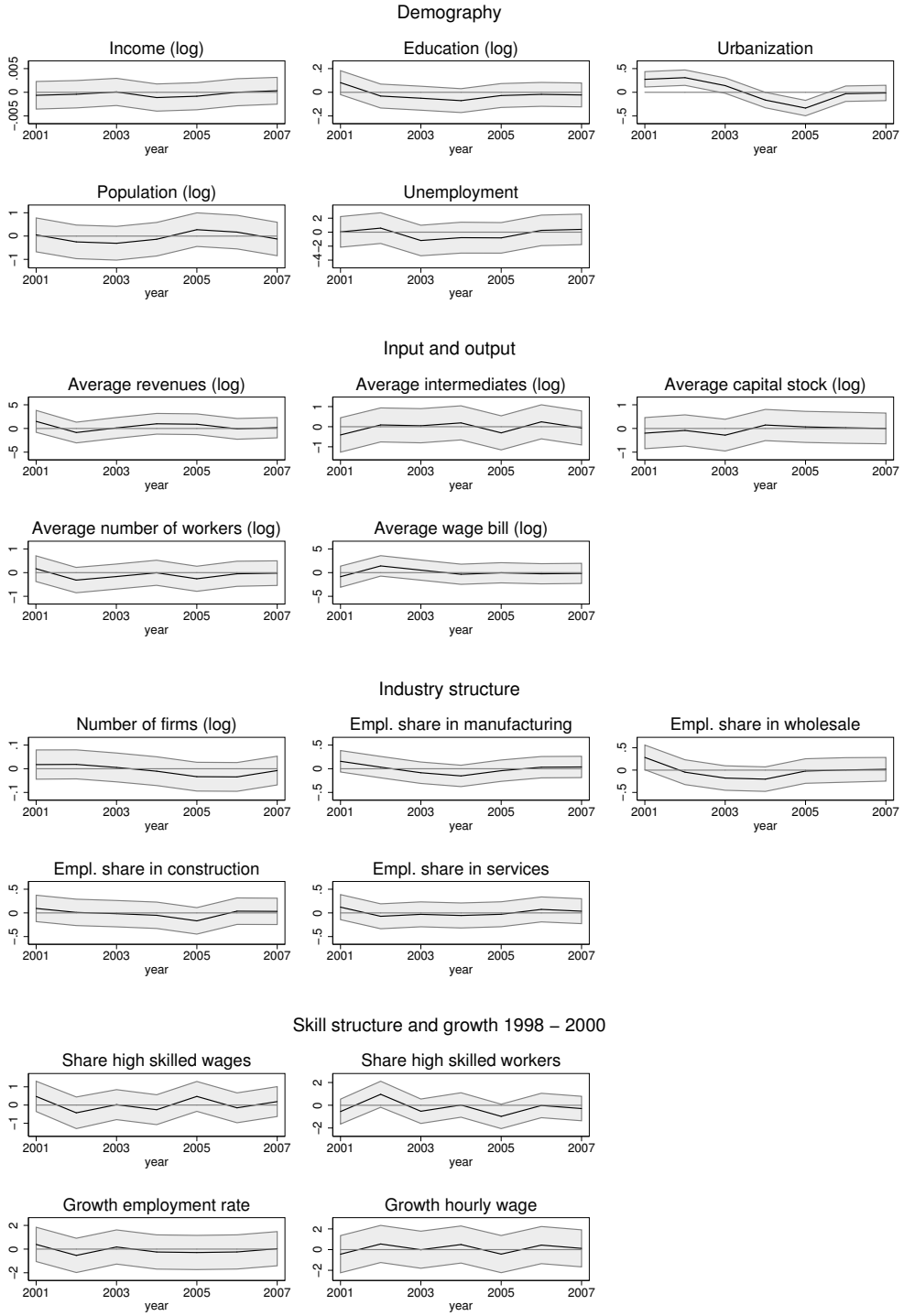
Note: The figures compare the weighted survey sample of joint-stock firms to the population of joint-stock firms. Detailed descriptions of the variables are given in Appendix Table A.1.

Figure A.3: Time trends in key firm variables



Note: The figures compare the weighted survey sample of joint-stock firms to the population of joint-stock firms. Detailed descriptions of the variables are given in Appendix Table A.1.

Figure A.4: Timing of broadband expansion and baseline covariates



Note: This figure report estimates from equation (2) of the vector ψ_t for every t (and the associated 95 % confidence intervals).

Table A.1: Variable definitions

Variable	Description
Firm accounts	Source: The Account Statistics.
Revenues	Total sales by a firm in year t.
Intermediates	Procurement of materials and intermediate inputs of a firm in year t.
Capital	Value of total fixed assets of a firm in year t.
Value added	Sales minus intermediates of a firm in year t.
Industry	4-digit code classifying a firm's main activity in year t according to the Nomenclature of Economic Activities (NACE2002) system.
Municipality	4-digit code for the municipality in which a firm is located in year t.
Exports	Total value of exported goods of a firm in year t.
Imports	Total value of imported goods of a firm in year t.
Internet variables	Source: The community survey on ICT in firms
Broadband	Dummy variable for whether a firm has adopted broadband internet (speed at or above 256 kilobits per second) in year t.
Revenues from online orders	Dummy variable for whether at least part of a firm's total revenues comes from online orders in year t.
Share of workers using a PC	Share of workers that use a PC in a firm in year t.
Employees	Source: Register of Employers and Employees and the Wage Statistics Survey.
Annual wages	Annual pre-tax wages in year t
Employment status	Dummy variable for whether annual wages exceed the substantial gainful activity threshold in year t (USD 6,850 in 2001), which defines employment in the Social Security System.
Hourly wages	Hourly pre-tax wage per October in year t.
Occupation	4-digit occupation code of a job in year t.
Individual characteristics	Source: National Education Database and Central Population Register.
Education level	Years of schooling.
Municipality	Municipality of residence in year t.
Age	The age of a worker in year t.
Potential experience	Age in year t - years of schooling - 7
Gender	The gender of a worker.

Variable	Description
Internet availability	Source: Norwegian Ministry of Government Administration.
Availability rate	Fraction of households in year t in a given municipality for which broadband internet is available, independently of whether they take it up.
Demographic controls	Source: Central Population Register.
Urbanization	Population share living in densely populated area in a given municipality in year t.
Income	Average annual disposable income across individuals aged 16–59 years in a given municipality in year t.
Education	Average years of schooling across individuals aged 16–59 in a given municipality in year t.
Unemployment	Unemployment rate among individuals aged 16–59 in a given municipality in year t.
Industry and firm controls	Source: The Account Statistics and Register of Employers and Employees.
Share of skilled workers	Share of employed workers with a college degree in a given municipality in year t.
Share of total wages to skilled workers	Share of the total wage bill paid to workers with a college degree in a given municipality in year t.
Share of employment by industry	Share of workers in the manufacturing/wholesale/service industry in a given municipality in year t.
Average input levels	Average level of capital stock/value added/number of workers/wages paid/revenues across firms in a given municipality in year t.
Growth in employment rate 1998-2000	Change from 1998 to 2000 in the average employment rate of workers aged 18-67 in a given municipality.
Growth in hourly wage 1998-2000	Proportional change from 1998 to 2000 in the average hourly wage of workers aged 18-67 in a given municipality.

Table A.2: Production function estimates

Dependent variable:	Log value added			
	OLS		LP	
	Population	Weighted Survey	Population	Weighted Survey
	(1)	(2)	(3)	(4)
Panel A: 1 skill category				
Intercept	1.971***	1.841***	4.172***	3.187***
	(0.0408)	(0.0851)	(0.0932)	(0.193)
Log capital	0.0780***	0.0821***	0.104***	0.164***
	(0.00297)	(0.00497)	(0.019)	(0.019)
Log labor	0.845***	0.856***	0.652***	0.676***
	(0.00351)	(0.00733)	(0.00480)	(0.00934)
Panel B: 2 skill categories				
Intercept	3.461***	3.380***	5.887***	4.695***
	(0.0455)	(0.0984)	(0.110)	(0.207)
Log capital	0.0990***	0.106***	0.107***	0.194***
	(0.00399)	(0.00599)	(0.022)	(0.019)
Log unskilled	0.558***	0.570***	0.410***	0.429***
	(0.0136)	(0.0141)	(0.0108)	(0.0143)
Log skilled	0.198***	0.194***	0.138***	0.135***
	(0.0115)	(0.0127)	(0.00913)	(0.0105)
Firm-year observations	149,676	16,744	149,676	16,744

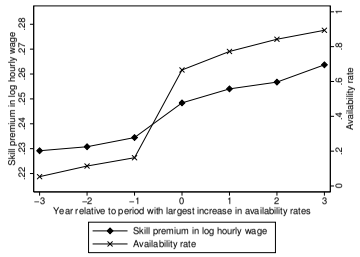
* p < 0.10, ** < 0.05, *** p < 0.01.

Note: The table reports estimates of Cobb-Douglas production functions, using the population of joint-stock firms over the period 2001-2007. The dependent variable is the log value added in a given year. Columns 2 and 4 restrict the sample to the survey sample. Sampling weights are used in columns 2 and 4 to ensure representative results for the population of joint-stock firms. (Un)Skilled comprises workers with(out) a college degree. All regressions include fixed effects for year, municipality and industry. The standard errors are clustered at the municipality level and robust to heteroskedasticity.

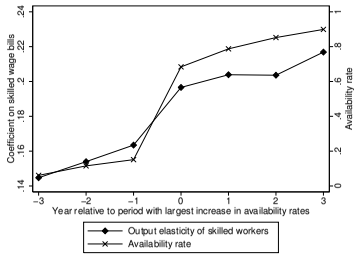
Appendix B: Specification checks and additional results

Figure B.1: Output elasticites and skill premiums conditional on year and municipality fixed effects with varying window, pre and post the largest increase in availability rates (period 0)

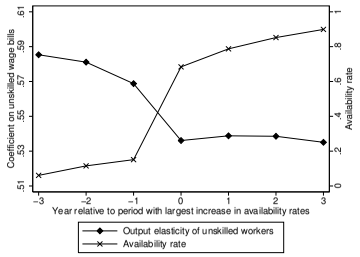
(a) Return to Skill: Hourly wage.



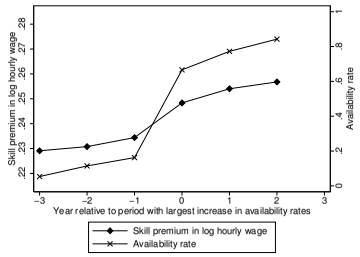
(b) Output elasticity: Skilled labor.



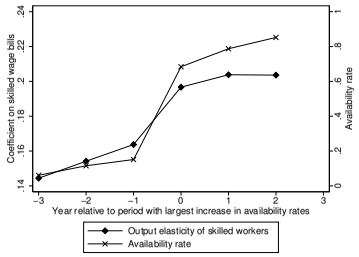
(c) Output elasticity: Unskilled labor.



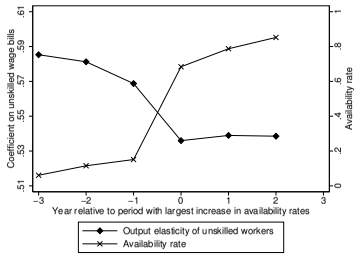
(d) Return to Skill: Hourly wage.



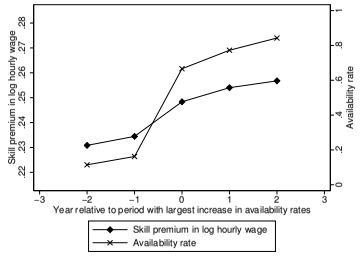
(e) Output elasticity: Skilled labor.



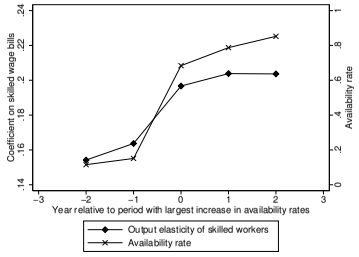
(f) Output elasticity: Unskilled labor.



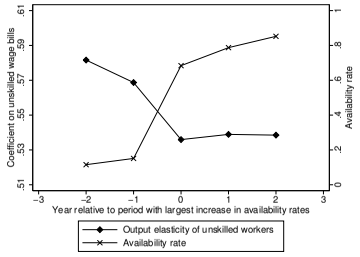
(g) Return to Skill: Hourly wage.



(h) Output elasticity: Skilled labor.



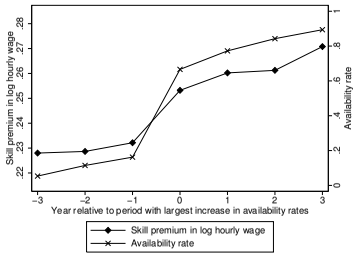
(i) Output elasticity: Unskilled labor.



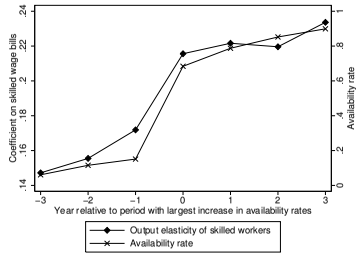
Note: Period zero represents the year with the strongest growth in availability rates in a given period. In each period, we estimate Cobb-Douglas production functions and wage regressions with year and municipality fixed effects. Graphs (b), (c), (e), (f), (h), and (i) report period-specific OLS estimates of the output elasticity of skilled and unskilled labor. Graphs (b), (d), and (g) report period-specific OLS estimates of log hourly wage on a dummy for skilled and controls for gender and potential experience.

Figure B.2: Output elasticites and skill premiums with varying window, pre and post the largest increase in availability rates (period 0)

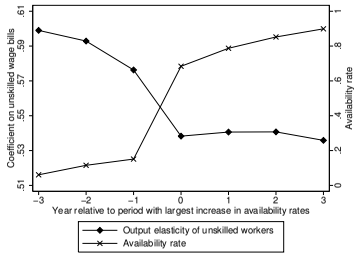
(a) Return to Skill: Hourly wage.



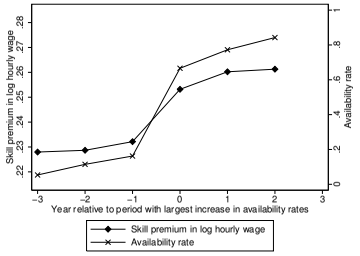
(b) Output elasticity: Skilled labor.



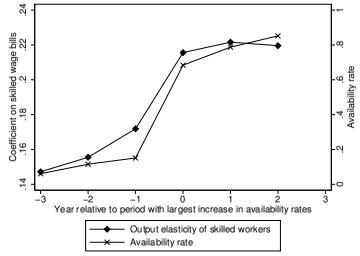
(c) Output elasticity: Unskilled labor.



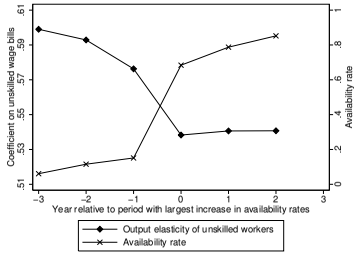
(d) Return to Skill: Hourly wage.



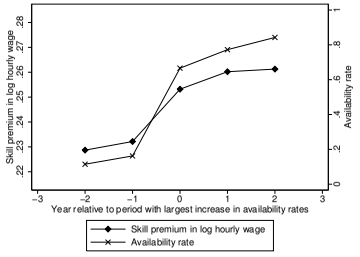
(e) Output elasticity: Skilled labor.



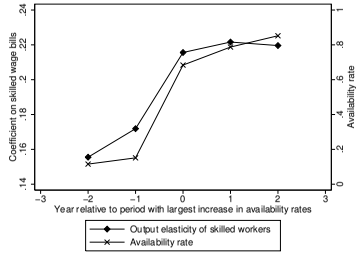
(f) Output elasticity: Unskilled labor.



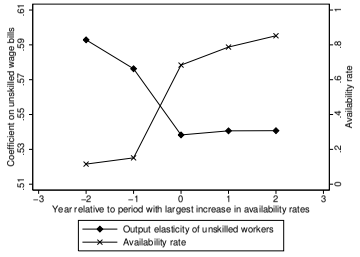
(g) Return to Skill: Hourly wage.



(h) Output elasticity: Skilled labor.

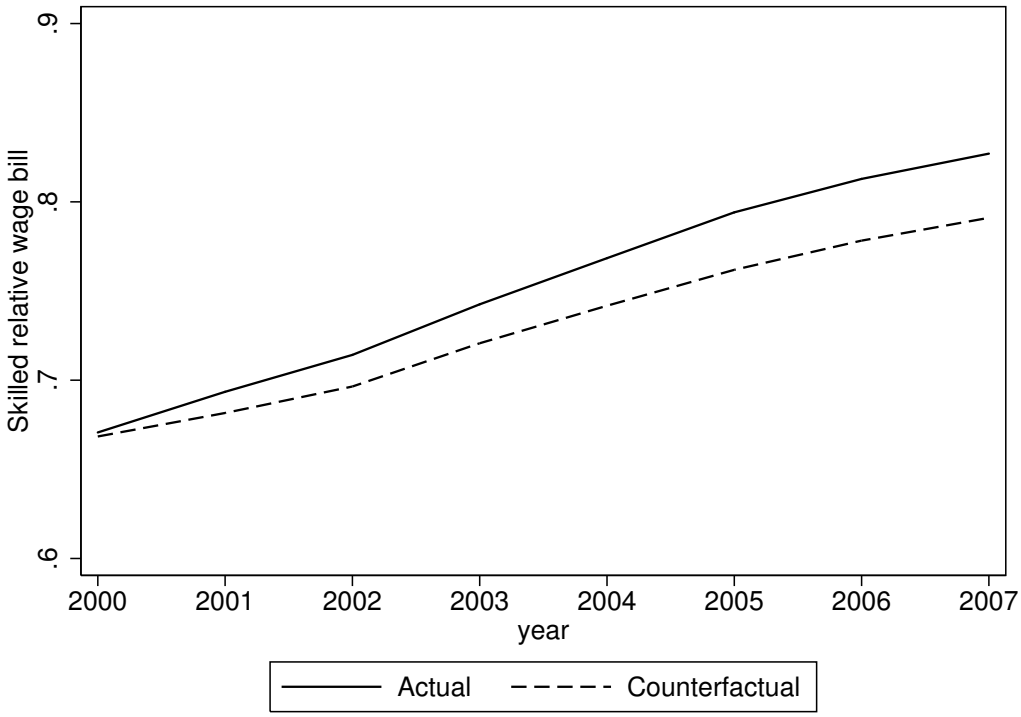


(i) Output elasticity: Unskilled labor.



Note: Period zero represents the year with the strongest growth in availability rates in a given period. In each period, we estimate Cobb-Douglas production functions and wage regressions. Graphs (b), (c), (e), (f), (h), and (i) report period-specific OLS estimates of the output elasticity of skilled and unskilled labor. Graphs (b), (d), and (g) report period-specific OLS estimates of log hourly wage on a dummy for skilled and controls for gender and potential experience.

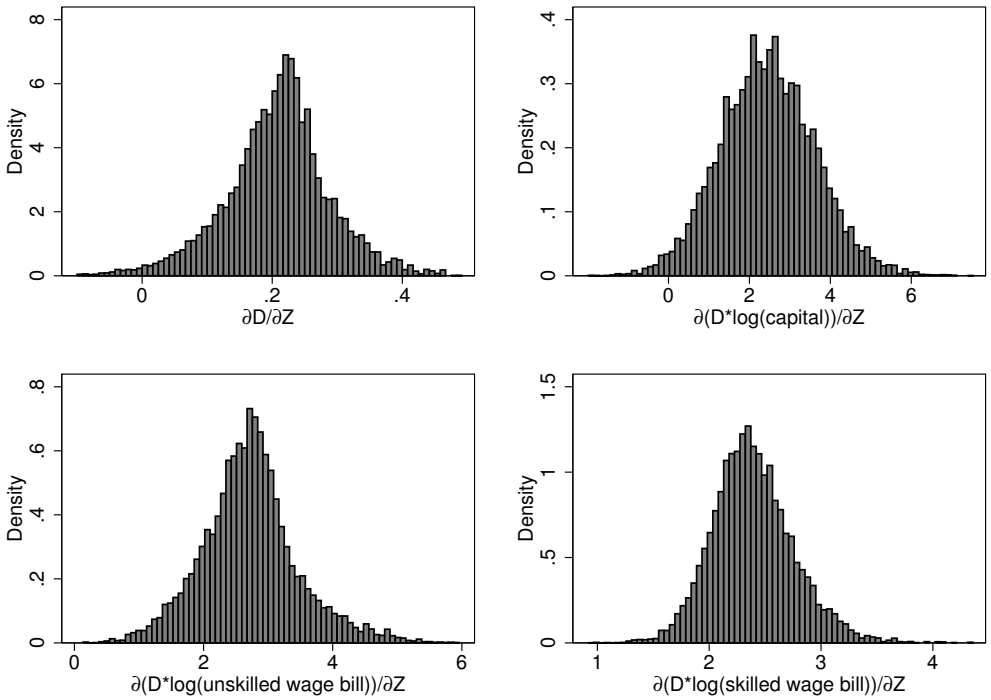
Figure B.3: Actual and counterfactual trends in relative wage bills.



Note: Solid line = actual outome. Dashed line = counterfactual outcome in the absence of broadband internet expansion.

The counterfactual outcome is measured as the actual outcome minus the predicted effect of broadband availability on the relative wage bill of skilled workers using the intention-to-treat estimates for hourly wages and employment in Table 3.

Figure B.4: Distribution of effects across firms of a marginal increase in broadband availability.



Note: These graphs use the first stage coefficients reported in Appendix Table B.12 to calculate the effects across firms of a marginal increase in broadband availability. We report one graph for each first stage dependent variable.

Table B.1: Specification checks for intention-to-treat effects on wages and employment

					Time interacted with covariates		
		Baseline	Covariates		Linear	Quadratic	Linear muncipality trends
		(1)	(2)	(3)	(4)	(5)	(6)
Panel A:	Unskilled	0.691***	0.691***	0.691***	0.697***	0.697***	0.696***
Employment rate		(0.00262)	(0.00143)	(0.00129)	(0.000967)	(0.00104)	(0.000927)
	Skilled	0.734***	0.733***	0.733***	0.739***	0.739***	0.738***
		(0.00480)	(0.00380)	(0.00363)	(0.00234)	(0.00231)	(0.00238)
	Availability×						
	Unskilled	0.000794	0.00170	0.00140	-0.00637***	-0.00679***	-0.00522**
		(0.00252)	(0.00122)	(0.00125)	(0.00245)	(0.00257)	(0.00245)
	Skilled	0.0208**	0.0217***	0.0214***	0.0141**	0.0138**	0.0153**
		(0.00920)	(0.00797)	(0.00775)	(0.00595)	(0.00588)	(0.00596)
Worker-year observations		20,327,515	20,276,208	20,276,208	20,233,574	20,233,574	20,276,208
Panel B:	Unskilled	2.939***	2.939***	2.940***	2.943***	2.943***	2.945***
Log hourly wage		(0.00455)	(0.00391)	(0.00356)	(0.00174)	(0.00181)	(0.00197)
	Skilled	3.169***	3.170***	3.170***	3.174***	3.174***	3.175***
		(0.00420)	(0.00361)	(0.00327)	(0.00174)	(0.00185)	(0.00190)
	Availability×						
	Unskilled	-0.00622	-0.00714*	-0.00790**	-0.0119***	-0.0118***	-0.0140***
		(0.00455)	(0.00376)	(0.00336)	(0.00190)	(0.00199)	(0.00208)
	Skilled	0.0178**	0.0168***	0.0161***	0.0116***	0.0117***	0.00924**
		(0.00720)	(0.00639)	(0.00592)	(0.00335)	(0.00347)	(0.00366)
Worker-year observations		8,759,388	8,739,814	8,739,814	8,724,567	8,724,567	8,739,814
Time-varying controls:							
Demographic			√	√	√	√	√
Industry				√	√	√	√

* p < 0.10, ** < 0.05, *** p < 0.01.

Note: Estimates are based on the model in equation (1), using worker-year observations over the period 2001-2007. Panel A considers the entire population of individuals between the ages of 18 and 67; the dependent variable is an employment dummy, taking the value of 1 if the individual is employed in a given year. Panel B considers the sample to workers aged 18-67 who are recorded in the wage statistics survey; the dependent variable is the log hourly wage in a given year. Column 2 adds demographic controls to the baseline model, including municipality-level information on average household income, mean years of schooling, share of population residing in a densely populated locality, size of population and level of unemployment. Column 3 also includes industry controls, consisting of municipality averages of revenues, intermediates, capital stock, number of workers and wage bills as well as employment share in manufacturing, employment share in wholesale, employment share in services, and shares of wages and workers by skill level. Columns 4 and 5 interact linear and quadratic time trends with baseline values of these covariates. Column 6 includes municipality-specific linear time trends. (Un)Skilled comprises workers with(out) a college degree. All regressions include fixed effects for year, municipality and industry and controls for gender, years of experience and years of experience squared. The standard errors are clustered at the municipality level and robust to heteroskedasticity.

Table B.2: Specification checks for intention-to-treat effects on output elasticities

Dependent variable:	Log value added					
				Time interacted with covariates		
	Baseline	Covariates		Linear	Quadratic	Linear municipality trends
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	3.880***	3.901***	3.876***	3.898***	3.896***	3.899***
	(0.0965)	(0.0964)	(0.0945)	(0.0961)	(0.0962)	(0.0967)
Log capital	0.100***	0.101***	0.101***	0.101***	0.101***	0.101***
	(0.00495)	(0.00499)	(0.00486)	(0.00495)	(0.00496)	(0.00501)
Log unskilled	0.576***	0.575***	0.577***	0.576***	0.576***	0.576***
	(0.0116)	(0.0116)	(0.0116)	(0.0117)	(0.0117)	(0.0119)
Log skilled	0.136***	0.136***	0.135***	0.135***	0.135***	0.134***
	(0.00678)	(0.00680)	(0.00670)	(0.00675)	(0.00676)	(0.00689)
Availability ×						
Intercept	-0.500***	-0.525***	-0.498***	-0.524***	-0.522***	-0.527***
	(0.111)	(0.111)	(0.107)	(0.110)	(0.110)	(0.112)
Log capital	-0.00169	-0.00232	-0.00249	-0.00282	-0.00287	-0.00284
	(0.00750)	(0.00752)	(0.00736)	(0.00749)	(0.00749)	(0.00757)
Log unskilled	-0.0226	-0.0216	-0.0238	-0.0231	-0.0232	-0.0232
	(0.0234)	(0.0234)	(0.0232)	(0.0234)	(0.0235)	(0.0238)
Log skilled	0.0755***	0.0761***	0.0766***	0.0774***	0.0776***	0.0781***
	(0.0166)	(0.0167)	(0.0166)	(0.0167)	(0.0167)	(0.0169)
Firm-year observations	149,676	149,610	149,610	149,482	149,482	149,610
Time-varying controls						
Demographic		√	√	√	√	√
Industry			√	√	√	√

* p < 0.10, ** < 0.05, *** p < 0.01.

Note: Estimates are based on the model in equation (1), using the population of joint-stock firms over the period 2001-2007. The dependent variable is the log value added of a firm in a given year. Column 2 adds demographic controls to the baseline model, including municipality-level information on average household income, mean years of schooling, share of population residing in a densely populated locality, size of population and level of unemployment. Column 3 also includes industry controls, consisting of municipality averages of revenues, intermediates, capital stock, number of workers and wage bills as well as employment share in manufacturing, employment share in wholesale, employment share in services, and shares of wages and workers by skill level. Columns 4 and 5 interact linear and quadratic time trends with baseline values of these covariates. Column 6 includes municipality-specific linear time trends. (Un)Skilled comprises workers with(out) a college degree. All regressions include fixed effects for year, municipality and industry. The standard errors are clustered at the municipality level and robust to heteroskedasticity.

Table B.3: Intention-to-treat effects on wages and employment: Alternative clustering and definiton of local labor markets

Dependent variable:	Log hourly wage			Employment		
	Baseline	Cluster at region	Regional level	Baseline	Cluster at region	Regional level
	(1)	(2)	(3)	(4)	(5)	(6)
Unskilled	2.939***	2.939***	2.926***	0.691***	0.691***	0.686***
	(0.00455)	(0.00441)	(0.00895)	(0.00262)	(0.00268)	(0.00736)
Skilled	3.169***	3.169***	3.160***	0.734***	0.734***	0.730***
	(0.00420)	(0.00287)	(0.00618)	(0.00480)	(0.00534)	(0.00857)
Availability×						
Unskilled	-0.00622	-0.00622*	0.00724	0.000794	0.000794	0.00773
	(0.00455)	(0.00326)	(0.00833)	(0.00252)	(0.00278)	(0.00887)
Skilled	0.0178**	0.0178***	0.0327***	0.0208**	0.0208**	0.0255*
	(0.00720)	(0.00660)	(0.0120)	(0.00920)	(0.00881)	(0.0130)
Worker-year observations	8,759,388	8,759,388	8,759,388	20,327,515	20,327,515	20,327,515

* p < 0.10, ** < 0.05, *** p < 0.01.

Note: Estimates are based on the model in equation (1), using worker-year observations over the period 2001-2007. Columns 1-3 consider the sample to workers aged 18-67 who are recorded in the wage statistics survey; the dependent variable is the log hourly wage in a given year. Columns 4-6 consider the entire population of individuals between the ages of 18 and 67; the dependent variable is an employment dummy, taking the value of 1 if the individual is employed in a given year. The standard errors in columns 1 and 4 are clustered at the municipality level and in columns 2-3 and 5-6 at the regional level (see Bhuller, 2009). In columns 3 and 6 we measure availability rates as the average availability rate at the regional level. (Un)Skilled comprises workers with(out) a college degree. All regressions include fixed effects for year, municipality and industry and controls for gender, years of experience and years of experience squared.

Table B.4: Intention-to-treat effects on output elasticities: Alternative clustering and definiton of local labor markets

Dependent variable:	Log value added		
	Baseline	Cluster at region	Regional level
	(1)	(2)	(3)
Intercept	3.880***	3.880***	3.952***
	(0.0965)	(0.113)	(0.148)
Log capital	0.100***	0.100***	0.104***
	(0.00495)	(0.00506)	(0.00830)
Log unskilled	0.576***	0.576***	0.563***
	(0.0116)	(0.0129)	(0.0157)
Log skilled	0.136***	0.136***	0.137***
	(0.00678)	(0.00668)	(0.00816)
Availability ×			
Intercept	-0.500***	-0.500***	-0.682***
	(0.111)	(0.0944)	(0.146)
Log capital	-0.00169	-0.00169	-0.00977
	(0.00750)	(0.00681)	(0.00822)
Log unskilled	-0.0226	-0.0226	-0.00122
	(0.0234)	(0.0237)	(0.0235)
Log skilled	0.0755***	0.0755***	0.0802***
	(0.0166)	(0.0197)	(0.0204)
Firm-year observations	149,676	149,676	149,676

* p < 0.10, ** < 0.05, *** p < 0.01.

Note: Estimates are based on the model in equation (1), using the population of joint-stock firms over the period 2001-2007. The dependent variable is the log value added in a given year. The standard errors in columns 1 are clustered at the municipality level and in columns 2 and 3 at the regional level (see Bhuller, 2009). In column 3 we measure availability rates at the regional level. (Un)Skilled comprises workers with(out) a college degree. All regressions include fixed effects for year, municipality and industry.

Table B.5: Intention-to-treat effects on output elasticities: LP approach and pre roll-out inputs

Dependent variable:	Log value added			
	Full sample		Firms observed in 2001	
	Baseline	LP	Baseline	Pre roll-out inputs
	(1)	(2)	(3)	(4)
Intercept	3.880***	5.936***	3.624***	3.730***
	(0.0965)	(0.156)	(0.105)	(0.123)
Log capital	0.100***	0.107***	0.0995***	0.0992***
	(0.00495)	(0.028)	(0.00507)	(0.00560)
Log unskilled	0.576***	0.444***	0.586***	0.589***
	(0.0116)	(0.0107)	(0.0119)	(0.0139)
Log skilled	0.136***	0.0948***	0.145***	0.148***
	(0.00678)	(0.00528)	(0.00650)	(0.00681)
Availability ×				
Intercept	-0.500***	-0.0341	-0.814***	0.00954
	(0.111)	(0.0999)	(0.126)	(0.121)
Log capital	-0.00169	-0.00193	-0.0126*	0.0124
	(0.00750)	(0.00573)	(0.00696)	(0.00901)
Log unskilled	-0.0226	-0.0412**	0.00818	-0.0548**
	(0.0234)	(0.0196)	(0.0247)	(0.0216)
Log skilled	0.0755***	0.0531***	0.0831***	0.0501***
	(0.0166)	(0.0126)	(0.0167)	(0.0153)
Firm-year observations	149,676	149,676	100,105	100,105

* p < 0.10, ** < 0.05, *** p < 0.01.

Note: Estimates are based on the model in equation (1). The dependent variable is the log value added of a firm in a given year. Columns 1 and 2 consider the population of joint-stock firms over the period 2001-2007 while columns 3 and 4 restrict the sample to joint-stock firms which are observed in 2001. (Un)Skilled comprises workers with(out) a college degree. In column 2 we apply the Levinsohn Petrin method. In column 4 we keep inputs fixed at the 2001 level. All regressions include fixed effects for year, municipality and industry. The standard errors are clustered at the municipality level and robust to heteroskedasticity.

Table B.6: Intention-to-treat effects on output elasticities in tradable and non-tradable sectors

Dependent variable:	Log value added				
		Above/below median			
	Baseline	Trade/Revenues		Geographic concentration	
		High	Low	High	Low
	(1)	(2)	(3)	(4)	(5)
Intercept	3.880***	3.492***	4.407***	3.751***	3.973***
	(0.0965)	(0.0955)	(0.106)	(0.120)	(0.0958)
Log capital	0.100***	0.106***	0.0923***	0.112***	0.0945***
	(0.00495)	(0.00692)	(0.00702)	(0.00695)	(0.00693)
Log unskilled	0.576***	0.561***	0.576***	0.568***	0.580***
	(0.0116)	(0.0157)	(0.0145)	(0.0191)	(0.0123)
Log skilled	0.136***	0.185***	0.0893***	0.152***	0.124***
	(0.00678)	(0.0104)	(0.00881)	(0.0102)	(0.00721)
Availability ×					
Intercept	-0.500***	-0.268*	-0.915***	-0.428**	-0.515***
	(0.111)	(0.140)	(0.155)	(0.167)	(0.129)
Log capital	-0.00169	-0.0131	0.0102	-0.00412	-0.00601
	(0.00750)	(0.00866)	(0.00956)	(0.00941)	(0.00803)
Log unskilled	-0.0226	-0.0249	-0.00799	-0.0405	-0.00550
	(0.0234)	(0.0274)	(0.0223)	(0.0308)	(0.0189)
Log skilled	0.0755***	0.0659***	0.0883***	0.0886***	0.0635***
	(0.0166)	(0.0184)	(0.0189)	(0.0206)	(0.0122)
Mean of tradability measure		0.28	0.02	0.00024	0.00007
Firm-year observations	149,676	74,619	75,057	68,379	81,297

* p < 0.10, ** < 0.05, *** p < 0.01.

Note: Estimates are based on the model in equation (1). The dependent variable is the log value added of a firm in a given year. Column 1 considers the population of joint-stock firms over the period 2001-2007. We use two measures of tradability. In columns 2 and 3, we measure tradability in each 4-digit industry by dividing total levels of exports and imports by the value added of firms. In columns 4 and 5, we follow J. B. Jensen and L. G. Kletzer (2005) "Tradable Services: Understanding the Scope and Impact of Services Outsourcing", Institute for International Economics Working Paper 05-09, in measuring tradability by the geographic concentration of an industry, defined as the Herfindahl index of employment shares across municipalities in each 4-digit industry. For each measure, we divide the sample into two groups: Industries with values of tradability above or below the median in the baseline firm sample. (Un)Skilled comprises workers with(out) a college degree. (Un)Skilled comprises workers with(out) a college degree. All regressions include fixed effects for year, municipality and industry. The standard errors are clustered at the municipality level and robust to heteroskedasticity.

Table B.7: Intention-to-treat effects on wages in tradable and non-tradable sectors

Dependent variable:	Log hourly wage				
		Above/below median			
	Baseline	Trade/Revenues		Geographic concentration	
		High	Low	High	Low
	(1)	(2)	(3)	(4)	(5)
Dependent variable: Log hourly wage					
Unskilled	2.939***	2.990***	2.952***	2.969***	2.946***
	(0.00455)	(0.00484)	(0.00547)	(0.00550)	(0.00469)
Skilled	3.169***	3.313***	3.176***	3.262***	3.209***
	(0.00420)	(0.0116)	(0.0129)	(0.0138)	(0.0105)
Availability ×					
Unskilled	-0.00622	0.0117	-0.0175***	0.000515	0.000264
	(0.00455)	(0.00792)	(0.00469)	(0.00345)	(0.00279)
Skilled	0.0178**	0.0325**	0.0325***	0.0276**	0.0522***
	(0.00720)	(0.0146)	(0.00965)	(0.0133)	(0.0121)
Mean of tradability measure		0.37	0.02	0.00038	0.00007
Worker-year observations	8,759,388	1,317,240	1,195,229	2,554,923	1,306,320

* p < 0.10, ** < 0.05, *** p < 0.01.

Note: Estimates are based on the model in equation (1), using worker-year observations over the period 2001-2007. Column 1 considers the sample of workers aged 18-67 who are recorded in the wage statistics survey; the dependent variable is the log hourly wage in a given year. In columns 2-9 we restrict the sample to workers employed by one of the firms in our sample. We use two measures of tradability. In columns 2, 3, 6 and 7 we measure tradability in each 4-digit industry by dividing total levels of exports and imports by the value added of firms. In columns 4, 5, 8 and 9, we follow J. B. Jensen and L. G. Kletzer (2005) "Tradable Services: Understanding the Scope and Impact of Services Outsourcing", Institute for International Economics Working Paper 05-09, in measuring tradability by the geographic concentration of an industry, defined as the Herfindahl index of employment shares across municipalities in each 4-digit industry. For each measure, we divide the sample into two groups: Industries with values of tradability above or below the median in the baseline firm sample. (Un)Skilled comprises workers with(out) a college degree. All regressions include fixed effects for year, municipality and industry and controls for gender, years of experience and years of experience squared. The standard errors are clustered at the municipality level and robust to heteroskedasticity.

Table B.8: Intention-to-treat effects on E-commerce and computerization

	Estimate (1)	Dependent mean (2)
Panel A: E-commerce:		
Dep. variable: Receiving orders online	-0.00265	0.26
	(0.0319)	
Panel B: Technical upgrading		
Dep. variable: Share of workers using a PC	-0.00217	0.58
	(0.0228)	
Worker-year observations	16,744	16,744

* p < 0.10, ** < 0.05, *** p < 0.01.

Note: This table uses the survey sample of joint-stock firms over the period 2001-2007. The table reports the coefficient for availability rates from a regression of the specified dependent variable on availability rates and municipality, industry and year fixed effects. Sampling weights are used to ensure representative results for the population of joint-stock firms. The standard errors are clustered at the municipality level and robust to heteroskedasticity.

Table B.9: Intention-to-treat effects on output elasticities excluding telecom firms and IT consultancy companies

Dependent variable:	Log value added		
	Baseline	No telecom	No IT consultancy
Intercept	3.880***	3.876***	3.824***
	(0.0965)	(0.0957)	(0.0980)
Log capital	0.100***	0.101***	0.0991***
	(0.00495)	(0.00499)	(0.00508)
Log unskilled	0.576***	0.575***	0.585***
	(0.0116)	(0.0116)	(0.0119)
Log skilled	0.136***	0.137***	0.131***
	(0.00678)	(0.00678)	(0.00697)
Availability ×			
Intercept	-0.500***	-0.476***	-0.459***
	(0.111)	(0.106)	(0.106)
Log capital	-0.00169	-0.00380	-0.00426
	(0.00750)	(0.00747)	(0.00764)
Log unskilled	-0.0226	-0.0215	-0.0212
	(0.0234)	(0.0233)	(0.0228)
Log skilled	0.0755***	0.0743***	0.0734***
	(0.0166)	(0.0166)	(0.0160)
Firm-year observations	149,676	148,973	144,579

* p < 0.10, ** < 0.05, *** p < 0.01.

Note: Estimates are based on the model in equation (1). The dependent variable is the log value added of a firm in a given year. Column 1 considers the population of joint-stock firms over the period 2001-2007. Column 2 excludes telecom firms (NACE code 64), whereas column 3 excludes IT consultancy firms (NACE code 72). (Un)Skilled comprises workers with(out) a college degree. All regressions include fixed effects for year, municipality and industry. The standard errors are clustered at the municipality level and robust to heteroskedasticity.

Table B.10: Intention-to-treat effects on wages and employment excluding telecom firms and IT consultancy companies

		Baseline (1)	No telecom employees (2)	No computer workers (3)
Panel A:	Unskilled	0.691***	0.689***	0.690***
Employment rate		(0.00262)	(0.00255)	(0.00267)
	Skilled	0.734***	0.732***	0.732***
		(0.00480)	(0.00468)	(0.00485)
	Availability×			
	Unskilled	0.000794	0.000906	0.00108
		(0.00252)	(0.00243)	(0.00257)
	Skilled	0.0208**	0.0201**	0.0210**
		(0.00920)	(0.00909)	(0.00931)
Worker-year observations		20,327,515	20,023,434	20,247,710
Panel B:	Unskilled	2.939***	2.937***	2.936***
Log hourly wage		(0.00455)	(0.00472)	(0.00450)
	Skilled	3.169***	3.169***	3.166***
		(0.00420)	(0.00413)	(0.00423)
	Availability×			
	Unskilled	-0.00622	-0.00526	-0.00668
		(0.00455)	(0.00481)	(0.00444)
	Skilled	0.0178**	0.0171**	0.0186**
		(0.00720)	(0.00700)	(0.00736)
Worker-year observations		8,759,388	8,538,873	8,679,584

* p < 0.10, ** < 0.05, *** p < 0.01.

Note: Estimates are based on the model in equation (1), using worker-year observations over the period 2001-2007. Panel A considers the entire population of individuals between the ages of 18 and 67; the dependent variable is an employment dummy, taking the value of 1 if the individual is employed in a given year. Panel B considers the sample to workers aged 18-67 who are recorded in the wage statistics survey; the dependent variable is the log hourly wage in a given year. Column 2 excludes workers in telecom firms (NACE code 64), whereas column 3 excludes workers in IT consultancy firms (NACE code 72). (Un)Skilled comprises workers with(out) a college degree. All regressions include fixed effects for year, municipality and industry and controls for gender, years of experience and years of experience squared. The standard errors are clustered at the municipality level and robust to heteroskedasticity.

Table B.11: Placebo test: Wages

Dependent variable:	Log hourly wage	
	Baseline sample of firms	Always/never taker firms only
	(1)	(2)
Unskilled	2.939***	2.916***
	(0.00455)	(0.0105)
Skilled	3.169***	3.171***
	(0.00420)	(0.0125)
Availability×		
Unskilled	-0.00622	0.0139
	(0.00455)	(0.0146)
Skilled	0.0178**	0.0135
	(0.00720)	(0.0189)
Worker-year observations	8,759,388	99,124

Note: Estimates are based on the model in equation (1), using worker-year observations over the period 2001-2007. In column 1, we consider workers aged 18-67 who are recorded in the wage statistics surveys. Column 2 restricts the sample to workers in firms that have adopted broadband even when the household availability rate is zero (always takers) and workers in firms that have not adopted broadband even when the household availability rate is one (never takers). (Un)Skilled comprises workers with(out) a college degree. Low skilled comprises individuals without high school diploma and medium skilled consists of high school graduates (without a college degree). All regressions include fixed effects for year, municipality and industry and controls for gender, years of experience and years of experience squared. The standard errors are clustered at the municipality level and robust to heteroskedasticity.

Table B.12: First stage regressions

Dependent variable:	Internet	Internet × Log capital	Internet × Log unskilled	Internet × Log skilled
	(1)	(2)	(3)	(4)
Intercept	-0.906***	-15.45***	-15.74***	-13.34***
	(0.188)	(2.141)	(2.238)	(2.081)
Log capital	0.0142	0.354***	0.182	0.163
	(0.0100)	(0.123)	(0.123)	(0.105)
Log unskilled	0.0428*	0.646***	0.860***	0.598**
	(0.0219)	(0.246)	(0.253)	(0.233)
Log skilled	0.0665***	0.849***	0.863***	0.905***
	(0.0151)	(0.173)	(0.187)	(0.152)
Availability ×				
Intercept	0.919***	5.630**	5.034**	4.311*
	(0.215)	(2.431)	(2.535)	(2.342)
Log capital	-0.00392	0.603***	-0.0512	-0.0402
	(0.0110)	(0.135)	(0.135)	(0.116)
Log unskilled	-0.0197	-0.344	0.298	-0.283
	(0.0245)	(0.277)	(0.283)	(0.260)
Log skilled	-0.0375**	-0.515***	-0.492**	0.189
	(0.0166)	(0.187)	(0.203)	(0.163)
Firm-year observations	16,744	16,744	16,744	16,744
F-value (excl. instruments)	41.4	56.7	38.5	28.6

* p < 0.10, ** < 0.05, *** p < 0.01.

Note: Estimates are based on the first stage regressions in equation (4), using the survey sample of joint-stock firms over the period 2001-2007. Sampling weights are used to ensure representative results for the population of joint-stock firms. (Un)Skilled comprises workers with(out) a college degree. All regressions include fixed effects for year, municipality and industry. The standard errors are clustered at the municipality level and robust to heteroskedasticity.

Table B.13: Broadband adoption and technological change: Levinsohn-Petrin

Dependent variable:	Log value added	
	2 skills	3 skills
	(1)	(2)
Intercept	4.225***	4.911***
	(0.535)	(0.505)
Log capital	0.216***	0.236***
	(0.040)	(0.039)
Log unskilled	0.537***	
	(0.0372)	
Log low skilled		0.278***
		(0.0283)
Log medium skilled		0.202***
		(0.0236)
Log skilled	0.0328	0.0500**
	(0.0236)	(0.0241)
Availability ×		
Intercept	-0.0511	-0.300
	(0.497)	(0.404)
Log capital	0.0110	0.00531
	(0.0251)	(0.0255)
Log unskilled	-0.151***	
	(0.0530)	
Log low skilled		-0.0830*
		(0.0426)
Log medium skilled		-0.0188
		(0.0365)
Log skilled	0.161***	0.127***
	(0.0365)	(0.0360)
Firm-year observations	16,744	16,250

* p < 0.10, ** < 0.05, *** p < 0.01.

Note: This table uses the survey sample of joint-stock firms over the period 2001-2007. The dependent variable is the log value added of a firm. The estimates are based on the model in equations (3) and (4) using the Levinsohn Petrin method. Sampling weights are used to ensure representative results for the population of joint-stock firms. (Un)Skilled comprises workers with(out) a college degree. Low skilled comprises individuals without high school diploma and medium skilled consists of high school graduates (without a college degree). All regressions include fixed effect for year, municipality and industry. The standard errors are clustered at the municipality level and robust to heteroskedasticity.

Table B.14: Occupation types, wages and task intensities by skill category

Occupation groups:	Proportion of workers			Mean relative wage		
	Professional & Managerial:	Production/operators/ clerical/retail:	Transport/farm/ construction/services:	Professional & Managerial:	Production/operators/ clerical/retail:	Transport/farm/ construction/services:
Primary tasks:	Abstract intensive	Routine intensive	Manual intensive	Abstract intensive	Routine intensive	Manual intensive
	(1)	(2)	(3)	(5)	(6)	(7)
A. 2 skill categories						
Unskilled	0.13	0.47	0.40	0.20	-0.14	-0.13
Skilled	0.65	0.23	0.12	0.30	0.00	0.02
B. 3 skill categories						
Low skilled	0.11	0.49	0.39	0.16	-0.15	-0.17
Medium skilled	0.15	0.44	0.41	0.23	-0.13	-0.09
Skilled	0.65	0.23	0.12	0.30	0.00	-0.01

Note: We consider workers aged 18-67 over the years 2001-2007 who are recorded in the wage statistics surveys and for which we observe occupation code at the 4-digit level. The occupation codes are linked with measures of task intensity from the Dictionary of Occupational Title, as reported by Autor and Dorn (2013). In Panel A, (Un)Skilled comprises workers with(out) a college degree. In Panel B, Low skilled comprises individuals without high school diploma and medium skilled consists of high school graduates (without a college degree). Following Table 2 in Autor and Dorn (2013), we divide occupation categories into three groups according to the measures of task intensities. The abstract intensive group includes managers, professional, technical, finance and public safety occupations. The routine intensive group includes production, craft, machine operatoring, assembly, clerical and retail sales occupations. The manual intensive group includes transport, construction, mechanical, mining, farm and service occupations. Columns 1-3 show the occupation distribution by skill category. Columns 4-6 provide mean relative wages by type of occupation and skill category. The mean relative wage is defined as the average log hourly wage within each skill-occupation group relative to the overall sample mean.

Table B.15: Examples of workplace tasks

Task measures		
Routine	Abstract	Manual
Record-keeping	Forming/testing hypotheses	Picking/sorting
Calculation	Medical diagnosis	Repetitive assembly
Repetitive customer service (e.g., bank teller)	Legal writing	Janitorial services
	Persuading/selling	Truck driving
	Managing others	

Source: Autor, Levy and Murnane (2003).

Appendix C: Marginal productivity and wages

To compare the changes in the marginal product and wages, we rewrite the intention-to-treat model in equation (1) such that all variables are in levels. Abstracting from fixed effects and control variables, this equation corresponds to a Cobb-Douglas production function with total factor productivity term and exponents on factor inputs that potentially change with the availability of broadband internet:

$$Y_{imt} = e^{\alpha_0+\alpha_1 z_{mt}} K_{imt}^{\delta_{k0}+z_{mt}\delta_{k1}} \left(w_{imt}^U H_{imt}^U\right)^{\delta_{u0}+z_{mt}\delta_{u1}} \left(w_{imt}^S H_{imt}^S\right)^{\delta_{s0}+z_{mt}\delta_{s1}}, \tag{6}$$

where Y_{imt} represents value added of firm i in municipality m in period t, K_{imt} is the capital stock, and w_{imt}^U and w_{imt}^S (H_{imt}^U and H_{imt}^S) denote the hourly wage (hours worked) of unskilled and skilled workers. In terms of equation (1), $(\alpha_0, \delta_{k0}, \delta_{s0}, \delta_{u0})$ is the coefficient vector δ_0, and $(\alpha_1, \delta_{k1}, \delta_{s1}, \delta_{u1})$ is the coefficient vector δ_1.

The marginal product of one hour of skilled labor input (Λ_{imt}^S) is defined as

$$\Lambda_{imt}^S \equiv \frac{\partial Y_{imt}}{\partial H_{imt}^S} = (\delta_{s0} + z_{mt}\delta_{s1}) \frac{Y_{imt}}{H_{imt}^S}$$

where $\delta_{s0} + z\delta_{s1}$ denotes the output elasticity of skilled labor.

To measure the pass-through from the broadband induced change in the marginal product of an hour worked to hourly wages, we compare a situation with no broadband availability ($z_{mt} = 0$) to a situation with full availability ($z_{mt} = 1$). In particular, we use the intention-to-treat effects on output elasticites and wages to assess how increased availability of broadband affects the marginal productivity and hourly wage of skilled workers. The latter is given directly from the intention-to-treat effects on log wages. The proportional change in Λ_{imt}^S due to an increase in broadband availability can be decomposed into three parts:

$$\frac{1}{\Lambda_{imt}^S}\frac{\partial \Lambda_{imt}^S}{\partial z} = \frac{\delta_{s1}}{(\delta_{s0} + z\delta_{s1})} + \frac{1}{Y_{imt}}\frac{\partial Y_{imt}}{\partial z_{mt}} - \frac{1}{H_{imt}^S}\frac{\partial H_{imt}^S}{\partial z_{mt}}$$

Using the the intention-to-treat effects on the production function (reported in Table IV), we calculate the first term on the right hand side: the proportional change in the output elasticity. To compute the second term, we use the same coefficients to calculate the mean proportional increase in value added holding all input factors constant. The third term is the proportional change in the use of skilled labor. Since we observe employment rates but not hours of work, we assume that all skilled workers are working the same number of hours and use the intention-to-treat effects on employment of skilled workers. In all calculations, we use the data on firms and workers in the population of joint-stock firms.

Our calculations show that the proportional increase in hourly wages is 3.7 percent while the proportional increase in the marginal product of an hour worked by a skilled worker is 19.0 percent. This suggests a pass-through rate of 19.5 percent. When we perform the same calculation for unskilled labor, we find an even smaller pass-through of changes in the marginal product to wages.

Appendix D: Levinsohn Petrin approach

The system of equations given in (3) and (4) is used to estimate production functions where firms can change their technology by adopting broadband internet. To address the concern that the factor inputs in x_{imt} might be correlated with broadband adoption and unobserved productivity, we follow LP and take a more structural approach to address this threat to identification of the production function.

LP use a structural model of an optimizing firm to derive the conditions under which intermediate inputs can be used to proxy for unobserved productivity in the production function. The error term ε_{imt} in (3) is assumed to be additively separable in a transmitted component (ω_{imt}) and an i.i.d. component (χ_{imt}). The key difference between ω_{imt} and χ_{imt} is that the former is a state variable, and therefore impacts the firm's decision rule, while the latter has no impact on the firm's decision. The intermediate input demand function depends on the firm-specific state variables, ω_{imt} and capital (k_{imt}),

$$a_{imt} = g_t(\omega_{imt}, k_{imt}). \tag{7}$$

and it must be monotonic in ω for all relevant k.[1] The monotonicity condition for intermediate inputs means that conditional on capital, profit maximizing behavior must lead more productive firms to use more intermediate inputs.

The monotonicity allows $g_t(\omega_{imt}, k_{imt})$ to be inverted to yield ω as a function of intermediate inputs and capital, $\omega_{imt} = \omega_t(a_{imt}, k_{imt})$. By expressing the unobserved productivity variable ω_{imt} as a function of observables, we are able to control for ω_{imt} in the second stage equation:

$$y_{imt} = x'_{imt}\beta_0 + D_{imt}x'_{imt}\beta_1 + w'_{imt}\theta + \lambda_m + \tau_t + \omega_t\left(a_{imt}, k_{imt}\right) + \chi_{imt}. \tag{8}$$

where β_0 is a vector consisting of $(\alpha_0, \beta_{k0}, \beta_{s0}, \beta_{u0})$, namely the pre roll-out intercept and output elasticities of capital, skilled labor, and unskilled labor. The vector β_1 is a vector consisting of $(\alpha_1, \beta_{k1}, \beta_{s1}, \beta_{u1})$ and measures the change in the intercept and the interaction effects between the input factors and broadband adoption. As in Olley and Pakes (1996) and LP, we use a polynomial expansion in a and k to approximate $\omega_t(\cdot)$. By simultaneous estimation of the first stage equations in (4) and the second stage equation in (3), we obtain consistent estimates of β_{u0}, β_{s0}, β_{k1}, β_{u1}, β_{s1}, and $\Phi_t\left(a_{imt}, k_{imt}\right) = \beta_{k0}k_{imt} + \omega_t(a_{imt}, k_{imt})$.

While these output elasticities are sufficient to assess how broadband adoption affects labor productivity, we need to identify β_{k0} to recover the full shift in production technology. Because k_{imt} is collinear with the non-parametric function $\omega_t(a_{imt}, k_{imt})$, further assumptions are necessary.[2]

[1] For simplicity, we assume as Olley and Pakes (1996) and Levinsohn and Petrin (2003) that capital is the only state variable over which the firm has control, while intermediates, labor and broadband internet are viewed as non-dynamic input factors.

[2] β_{k1} is identified as the interaction of capital with D_{imt} provides independent variation. Note also that the intercept in the production function is not separately identified from the mean of $E\left[\omega_{imt}|\omega_{imt-1}\right]$ without some further restriction.

Assuming that ω_{imt} follows a first-order Markow process, we can write

$$\omega_{imt} = E\left[\omega_{imt}|\omega_{imt-1}\right] + \xi_{imt}.$$

This simply decomposes ω_{imt} into its conditional expectation at time $t-1$, $E\left[\omega_{imt}|\omega_{imt-1}\right]$, and a deviation from that expectation, ξ_{imt}. If the capital stock is pre-determined and current investment (which will react to productivity shocks) takes one period before it comes productive, it follows that

$$E\left[\xi_{imt}k_{imt}\right] = 0.$$

This is the moment which LP use to identify the capital coefficient. Roughly speaking, variation in k_{imt} conditional on ω_{imt-1} is the exogenous variation used for identification. To operationalize this approach in a GMM context, note that given a guess on the capital coefficient β_{k0}, we can rewrite unobserved productivity as

$$\omega_{imt}(\beta_{k0}) = \hat{\Phi}_{imt} - \beta_{k0}k_{imt}.$$

Given these $\omega_{imt}(\beta_{k0})$, we compute ξ_{imt} by non-parametrically regressing $\omega_{imt}(\beta_{k0})$'s on $\omega_{imt-1}(\beta_{k0})$'s and a constant term; we then form the residual

$$\xi_{imt}(\beta_{k0}) = \omega_{imt}(\beta_{k0}) - \hat{\Psi}\left(\omega_{imt-1}(\beta_{k0})\right)$$

where $\hat{\Psi}\left(\omega_{imt-1}(\beta_{k0})\right)$ are predicted values from the non-parametric regression.

The $\xi_{imt}(\beta_{k0})$'s are used to form a sample analogue to the above moment. i.e.

$$\frac{1}{T}\frac{1}{N}\sum_{t}\sum_{i}\xi_{imt}\left(\beta_{k0}\right)\cdot k_{imt}$$

where N denotes number of firms and T number of time periods. We estimate β_{k0} by minimizing the GMM criterion function

$$Q\left(\beta_{k0}\right) = \min_{\beta_{k0}}\left(\frac{1}{N}\frac{1}{T_{i1}}\sum_{i}\sum_{t=T_{i0}}^{T_{i1}}\xi_{imt}\left(\beta_{k0}\right)\cdot k_{imt}\right)^{2}$$

where i indexing firms and T_{i0} and T_{i1} index the second and last period in which firm i is observed.

Because our baseline sample is a repeated cross-section (rather than panel data), we adjust the above estimation procedure. Exploiting the random sampling of firms, we can identify β_{k0} from the moment

$$E\left[\bar{\xi}_{mt}, \bar{k}_{mt}\right] = 0.$$

where the municipality average of a variable is denoted by upper bar. By applying the above procedure to our panel data at the municipality level, we obtain the GMM criterion function

$$Q\left(\beta_{k0}\right) = \min_{\beta_{k0}}\left(\frac{1}{M}\frac{1}{T_{m1}}\sum_{m}\sum_{t=T_{m0}}^{T_{m1}}\sqrt{N_{mt}}\bar{\xi}_{mt}\left(\beta_{k0}\right)\cdot \bar{k}_{mt}\right)^{2}$$

where T_{m0} and T_{m1} index the second and last period municipality m is observed and N_{mt} is the number of firms in municipality m in period t. To obtain standard errors on β_{k0}, we use bootstrap while clustering by municipality.