MANCHESTER 1824
The University of Manchester

Modelling Emerging Technologies in the Green Goods Sector

Marianne Sensier
marianne.sensier@manchester.ac.uk

11/06/2014,
Summer School

ESRC Sustaining Innovation Project

Overview

- The Green Goods Project
- Defining and measuring the green goods sector
- Search based method
- UK firm profile in the green good sector
- Cross-section method and results
- Panel method and results
- Alternative Approaches

ESRC Sustaining Innovation Project

- We analyse the growth strategies of innovative SMEs in "green goods" industries in the UK, USA and China. Firms in these sectors are strategically important to economic and environmental rebalancing and have significant potential for sustainable growth.
- Our key questions include:

1. What differentiates the business strategies and technology pathways of successful enterprises?
2. How does regional clustering influence business strategies, technology pathways, and relationships between firms?
3. What are the contributions of policy-induced resources to SME growth?

ESRC SI Project: Motivation

- We investigate the drivers of growth in SMEs.

1. What is the role of Research and Development (R&D)? Griffith, et al (2006) AER.
2. How important is regulation from Governments, see Rennings and Rammer (2011), Ind & Innov.
3. Is the increasing internalisation of firms playing a role? Global or local connections?
4. Does family ownership or professional management contribute to growth?
5. Are early alliances important for the firm, see Mohr, et al (2013), ICC?

ESRC SI Project: Literature

- Coad and Rao (2011), J of Evoluntionary Economics, study a panel of US high-tech firms and find that:

1. The size of the firm effects the amount of R&D activity and growth patterns: innovative activity in large firms is more associated with employment growth.
2. Firms in different sectors behave differently (see also Rennings and Rammer, 2011, Industry and Innovation).
3. They consider that product innovations generally have a positive impact on employment, whilst the role of process innovations is more ambiguous (see also Hall et al, 2008, ICC).

ESRC SI Project: Literature

- Coad and Rao (2011), estimate a quantile regression with model:

$$GROWTH_{i,t} = \alpha + \beta_1 INN_{i,t} + \beta_2 CONTROL_{i,t-1} + y_t + \varepsilon_{i,t}$$

- The dependent variable, GROWTH, is the growth rate of employment.
- Where INN is an innovative index using Principal Components Analysis, creating an index from firm's patents and R&D expenditure.
- CONTROL includes lagged growth and lagged size dummies. y are the year dummies.

Introduction: Green Goods Sector

- The green goods sector (GGS) – comprises of companies in a range of industries that produce or market manufactured items that have environmental or natural resource benefits when used by other businesses, organizations or households.
- Green industries help address the challenges of environmental sustainability, but require coordination between government, industry and academia (Triple Helix literature, see Etzkowitz & Leydesdorff, 2000, Research Policy).

UK Project Database overview

- UK firms financial data from FAME, BvD (2012). Including firm turnover, number of employees, R&D spending and overseas turnover.
- Patents: total number of green patents received as noted in the Derwent World Patents Index, Thomson (2010).
- Publications: total number of published articles as recorded in the Web of Science database.
- Grants: total number of grants received by firms from Technology Strategy Board (TSB) database.

UK Firm's FAME Search

We restricted the search to include firms with:

- Incorporation date: from 01/01/1995 to 31/12/2007.
- Number of Employees: for the range 2002 to 2007, min=1, max=500.
- Mostly Manufacturing SIC but widened for some.
- We searched using key words for each sector in the firm's "Overview" and "Trade Description" sections, for example in the Air Pollution sector: (((air* contr*) OR (dust* contr*) OR (particular* contr*) OR (air* qual*)) AND (pollut*)).

UK Firms FAME Search Refined

- We arrived at 500 firms, then reduced this to 304 by firm relevance, see Shapira et al (2014).
- Final number of firms dropped to 296.
- The 5 most popular Manufacturing SIC codes include: Manufacture of electrical equipment (12.7%); Manufacture of machinery and equipment n.e.c. (11.3%); Manufacture of fabricated metal products, except machinery and equipment (11%); Manufacture of motor vehicles, trailers and semi-trailers (9.9%) and Manufacture of chemicals and chemical products (7.2%).
- SIC in Services 12% and Construction 3%.

FAME/Experian Employment in 2012

Employee Count:	Employment	Employment and Turnover
Between 1-10	56	25
Between 11-50	78	58
Between 51-250	117	103
More than 250	24	19
Missing data	21	0
Total	296	205

Source: FAME and additional data provided by Experian.

FAME Ownership Type in 2012

Ownership Category:	Employment	Employment and Turnover
Independent	139	91
Subsidiary of a larger UK company	49	32
Subsidiary of a larger foreign company	95	71
Holding company	12	11
Eco-innovation firm	1	0
Total	296	205

Source: FAME.

FAME Legal Form in 2012

Legal Form:	Employment	Employment and Turnover
Guarantee	1	1
Private Limited	268	181
Public AIM	18	16
Public, Not Quoted	7	5
Public, Quoted	2	2
Total	296	205

Source: FAME.

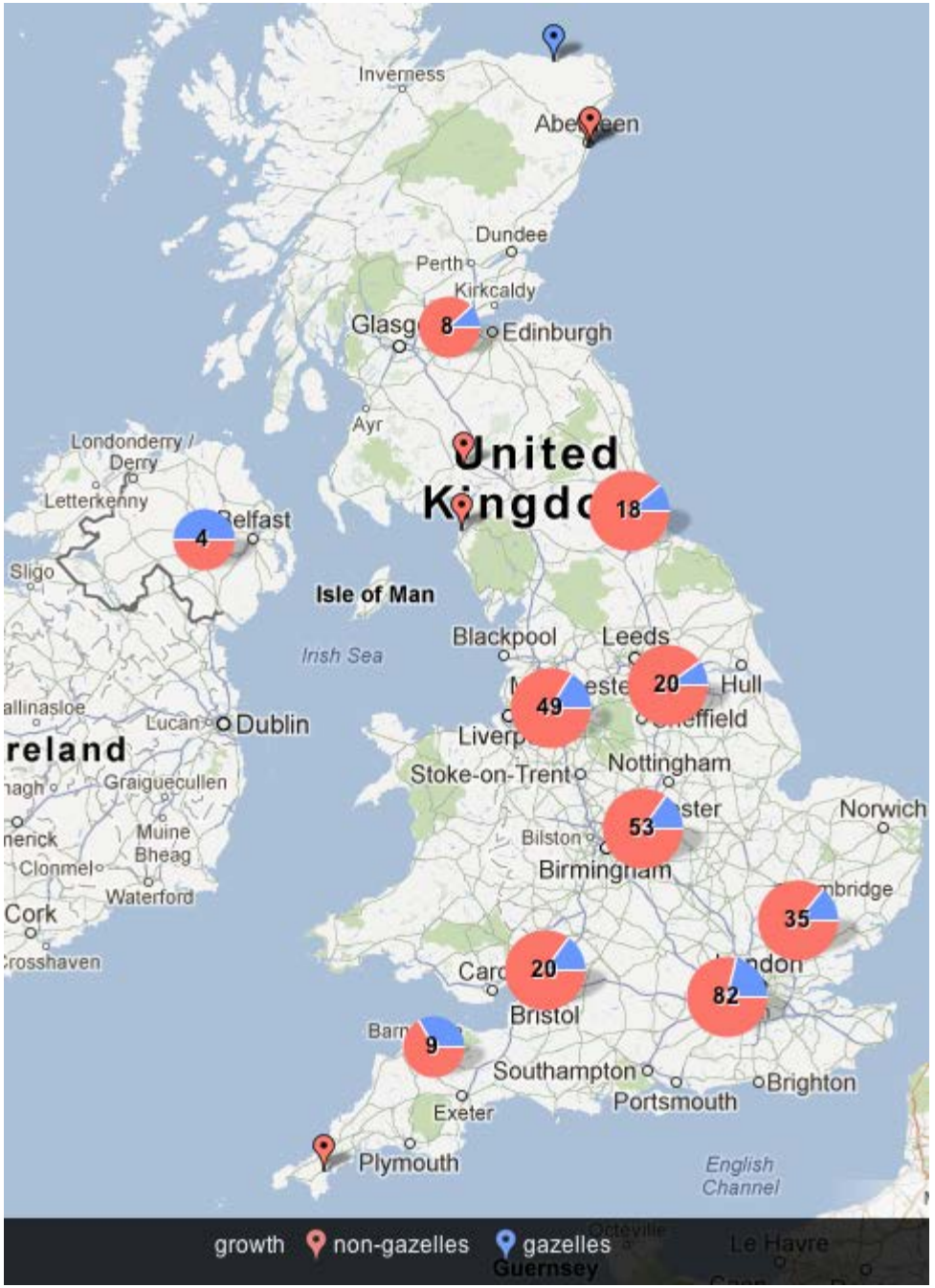
Sectors in 2012

Sector:	Employment	Employment and Turnover
Pollution control (E)	23	19
Building Technologies (LC)	73	63
Battery (LC)	13	10
Alternative vehicle/fuel (LC)	67	36
Renewables (R) and water treatment (E)	59	37
Miscellaneous	61	40
Total	296	205

Source: Noted by researchers. LCEGS code: E – Environment; LC – Low Carbon and R – Renewables.

UK Location by Region

NUTS 1:	Emp.	E&T
North East	16	11
North West	38	29
Yorks. & H.	23	19
East Midlands	25	15
West Midlands	32	21
Eastern	36	27
London	35	21
South East	42	27
South West	20	15
Wales	13	10
Scotland	12	7
N. Ireland	4	3



Web variable keywords searched

Web variable:	Employment and Turnover
Regulation (nreg)	Regulat*, law, ordinance, statue, decree, convention, guideline, rule, standard, directive, parliament
R&D Intensity (nrndins7)	lab, laboratory, research, development, R&D,researcher, scientist
Manufacturing Intensity (nmanuint)	manufactur produc assembly, *fabricat*, machined, machinery, machines, mold, prefab, process, construct, weld, engineer
Partnerships (npartner)	partner, stakeholder, distributor, collaborat, alliance, joint venture, agreement, supplier, parts manufacturer
Rest of World Cities (nrowcitiesins)	List of Cities

Source: Firms web-sites.

Dummy Variables for R&D

- As we have a very small number of firms with patents, publication, grants and R&D spending we combine these:

ResearchDummy = patents + publications + grants + R&D spending

- To include the effect of R&D keywords mentioned on the firm's web-site we also add in the R&D web variable:

webRnDdummy = ResearchDummy+nrndins7

Dummy Variable Indicators in 2012

Dummy Variables:	Total_296	Subset_205
Patents	2.7%	1.5%
Publications	0.3%	0
Grants	5.7%	6.8%
R&D spending	14.2%	19.5%
ResearchDummy	21%	25.4%
webRnDdummy	66.2%	87.3%
Overseas Turnover	35.1%	45.4%
Family Ownership	13.2%	13.7%

Source: Family Ownership data provided by Experian.

Any questions?

The Linear Probability Model (LPM): 1

- The multiple regression model is written as:

(1) $$y = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k + u$$

- The dependent variable, y, takes on two values:

(2) $$y = \begin{cases} 0 \text{ with prob } 1-p \\ 1 \text{ with prob } p \end{cases}$$

- We assume the zero conditional mean holds (does not depend on the explanatory variables):

(3) $$E(y|\mathbf{x}) = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k$$

- The probability of success is a linear function of the x_j, $P(y=1|\mathbf{x}) = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k$

The Linear Probability Model: 2

- The multiple linear regression model with a binary dependent variable is called the linear probability model (LPM) as the response probability is linear in the parameters β_j.
- In the LPM β_j measures the change in the probability of success when x_j changes, holding other factors fixed: $\Delta \mathrm{P}\left(y=1 \middle| \mathbf{x}\right) = \beta_j \Delta x_j$
- Our probability of success is when the growth of UK firm real turnover is positive.

Problems with the LPM

1. The predicted y is the predicted probability of success. A potential problem is that it can be outside the [0,1] interval.
2. When y is a binary variable, its variance, conditional of x is: $\text{Var}(y|\mathbf{x}) = p(\mathbf{x})[1 - p(\mathbf{x})]$

This means there must be heteroscedasticity in the LPM. This does not cause bias for the OLS estimates of β_j but the standard errors (s.e.) are not generally valid. This can be adjusted in standard econometric packages. For example in Stata 12 use vce(robust) for White-corrected s.e.

UK Firm's Growth: Real turnover

- We calculate the firm's real turnover by taking nominal turnover and dividing it by the Consumer Price Index (2005=100 base year):

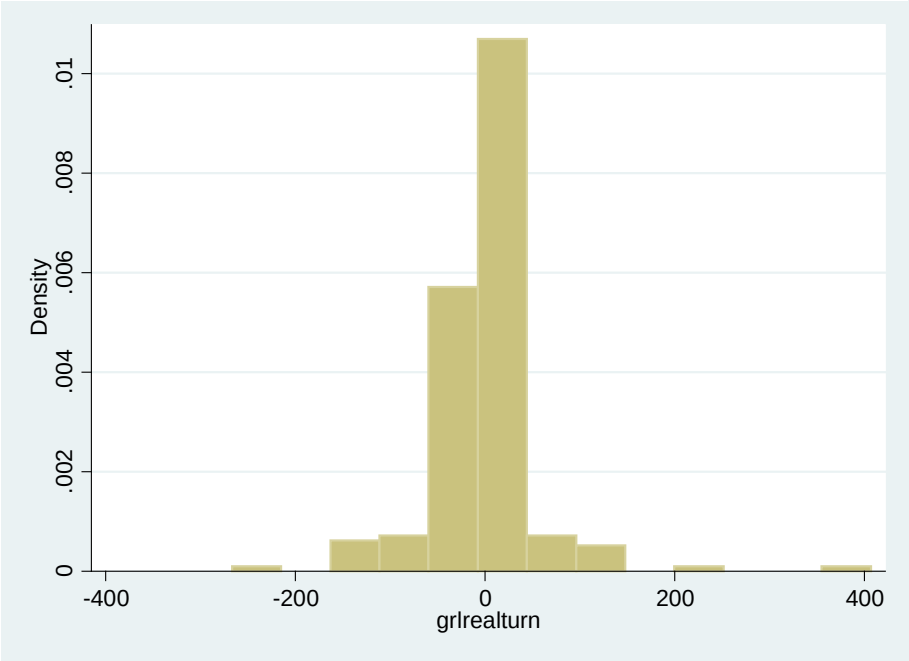
  Realturn = turnover/CPI

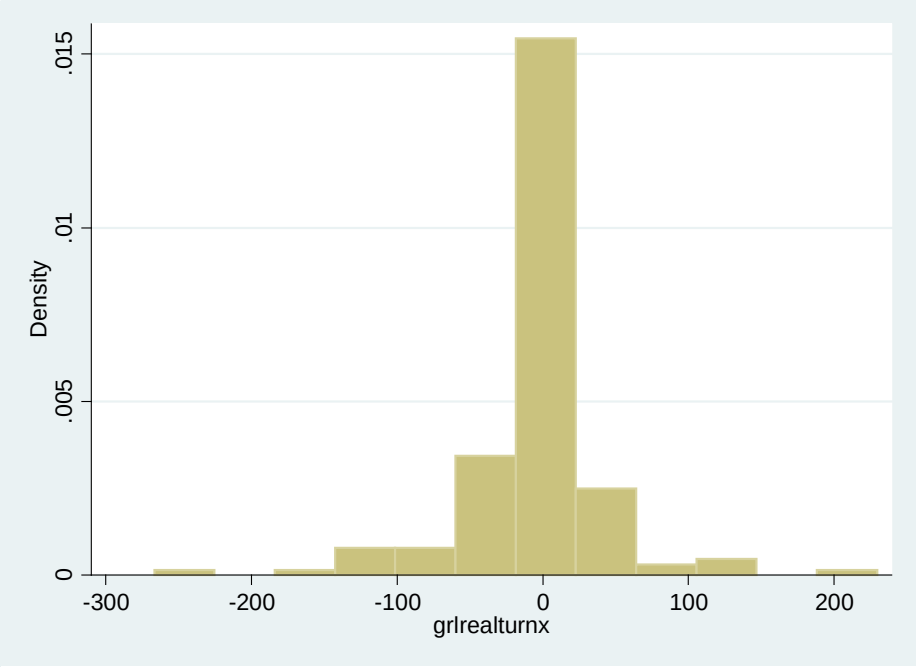
- The growth rate is calculated by taking the first difference of the natural log of the data x 100:

  $dRealturn = 100[\log(Realturn)_t - \log(Realturn)_{t-1}]$

- To analyse the distribution on growth in real turnover we plot the following histograms.

UK Firms Growth: Real turnover





Histograms for the growth of real turnover in 2012:

- LHS chart for all observations of growth in real firm turnover.
- RHS chart for observations that match with employment and web variables.

UK Firm's Employment

- As we have some missing observations for some firm's employment we fill in values some with a moving average smoother:

  Stata> tssmooth ma MA4emp=emp, window(3 1 0)

The smoother applied was by firm MA4emp=

(1/4)*[x(t-3) + x(t-2) + x(t-1) + 1*x(t)]; x(t)= emp

- We calculate the growth rate transformation of the data (not logs as much smaller numbers):

  gen dMA4emp=(MA4emp-l.MA4emp)/MA4emp

- We include employment as the lagged growth rate:

Stata> generate lagdMA4emp=l.dMA4emp

Cross Section Analysis Results

Methodological Choices: cross-section

- LPM cross-section model for firms in two years.
- Dependent variable is growth of real turnover in year 2012 compared to 2008 before recession hit.
- Web variables are standardised so the count of the keyword (webvar) is divided by the number of phrases on the web-site:

  nwebvar = (webvar/phrases)*100
- Dummy variables included for Regions (controls).
- Experian firm financial robustness variables checked are delphi and phMegascore.

UK Firm's Growth and Employment

Source	SS	df	MS	Number of obs =	241
				F(1, 239) =	2.35
Model	.564888658	1	.564888658	Prob > F =	0.1264
Residual	57.3936176	239	.240140659	R-squared =	0.0097
				Adj R-squared =	0.0056
Total	57.9585062	240	.241493776	Root MSE =	.49004

grtposdum12	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lagdMA4emp	.074774	.0487531	1.53	0.126	-.0212666	.1708147
_cons	.4067495	.0316883	12.84	0.000	.3443255	.4691736

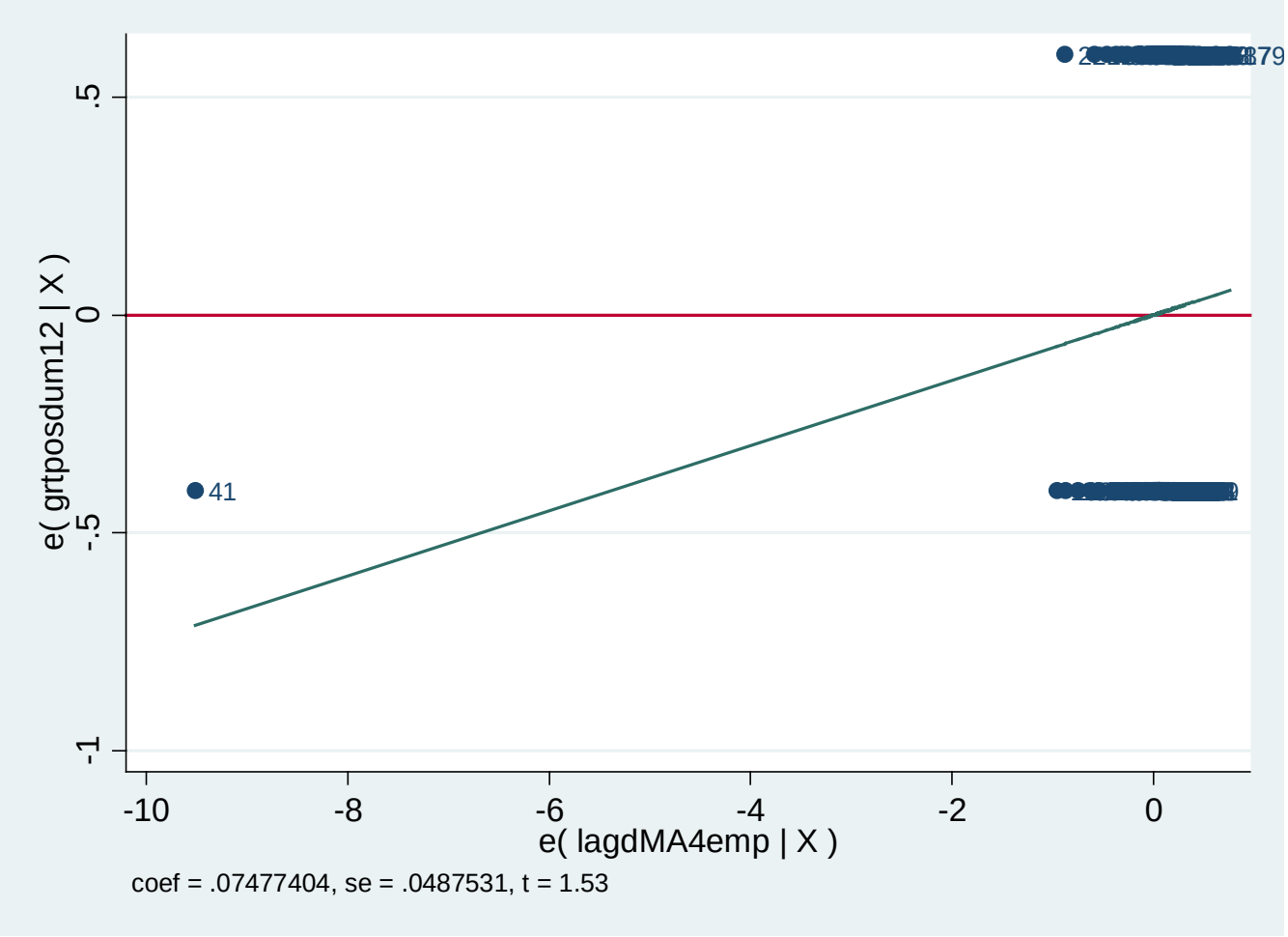
```
. avplot lagdMA4emp, mlabel(firm) yline(0)

. reg grtposdum12 lagdMA4empx
```

Source	SS	df	MS	Number of obs =	241
				F(1, 239) =	6.27
Model	1.48159891	1	1.48159891	Prob > F =	0.0129
Residual	56.4769073	239	.236305052	R-squared =	0.0256
				Adj R-squared =	0.0215
Total	57.9585062	240	.241493776	Root MSE =	.48611

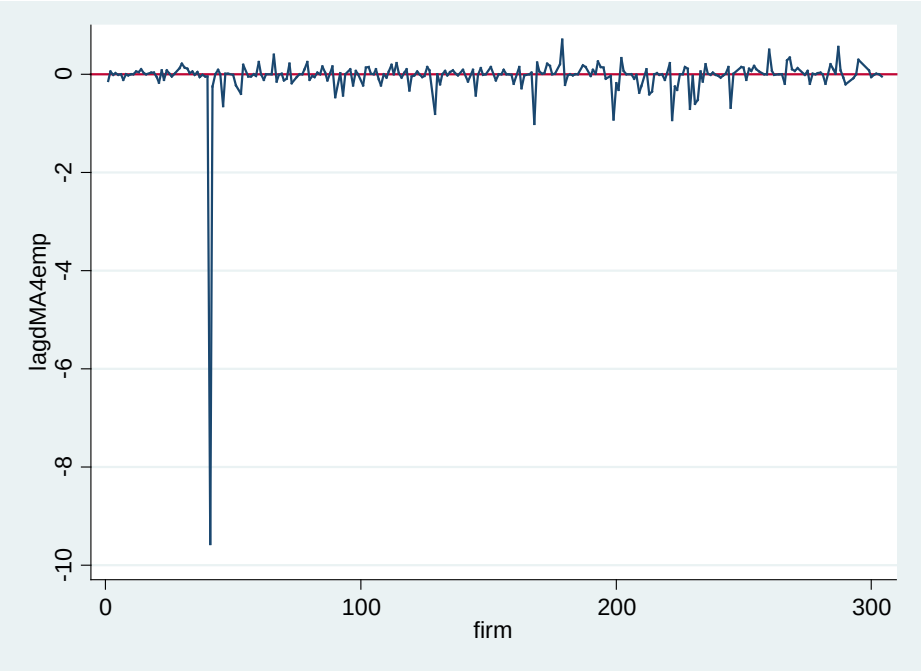
grtposdum12	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lagdMA4empx	.376786	.1504755	2.50	0.013	.0803583	.6732137
_cons	.4098843	.0314522	13.03	0.000	.3479254	.4718432

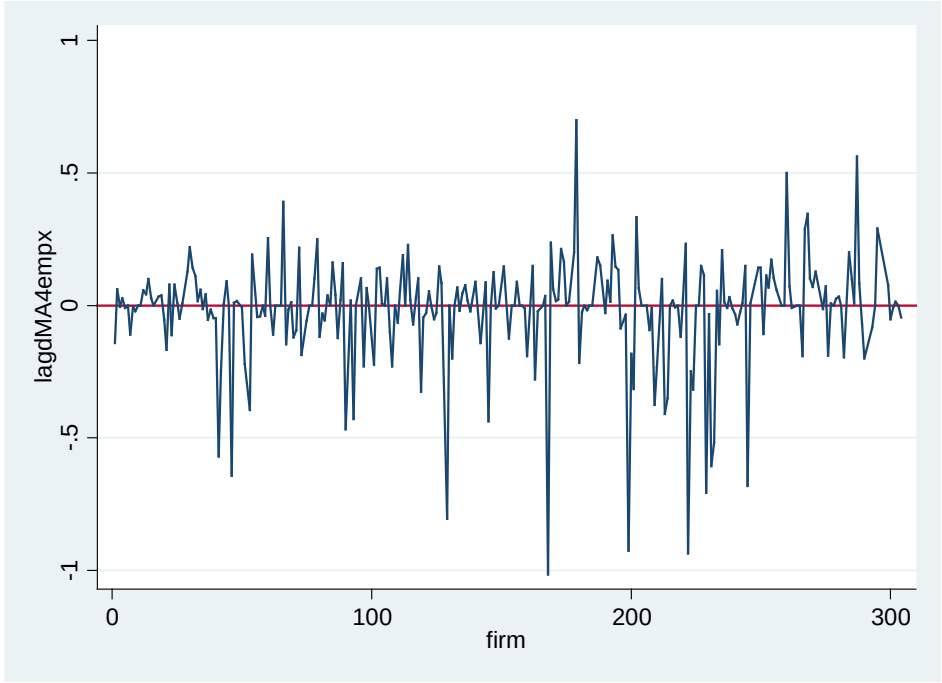
UK Firm's Employment: outliers

- We notice an outlier for firm employment:



UK Firm's Employment: outliers





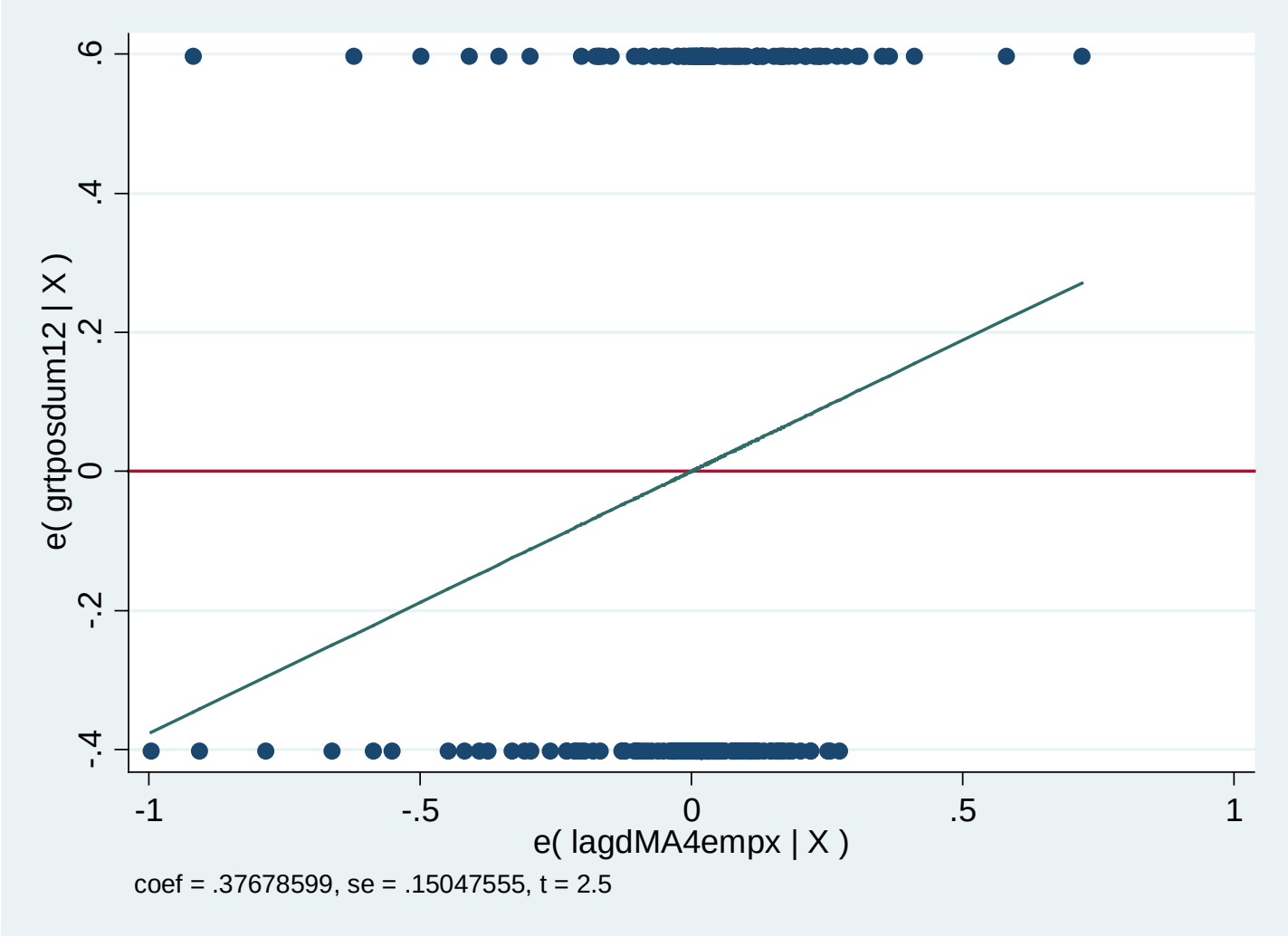
Firm's employment with the outlier on the left and with the outlier amended on the right.

Stata> twoway (line lagdMA4emp firm), yline(0) and

Stata> twoway (line lagdMA4empx firm), yline(0)

UK Firm's Employment: corrected

- Reducing the extreme obs. from-9.57 to -0.57:



LPM cross-section results for 2012

Linear regression

Number of obs = 205
F(16, 188) = 2.98
Prob > F = 0.0002
R-squared = 0.1418
Root MSE = .47079

grtposdum12	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
lagdMA4empx	.3479205	.1931445	1.80	0.073	-.0330884	.7289294
nreg	.0864376	.0557343	1.55	0.123	-.0235074	.1963825
nmanuint	.0238606	.0121388	1.97	0.051	-.0000852	.0478063
nrowcitiesins	.0135096	.0097902	1.38	0.169	-.0058032	.0328223
sizeSmall	.1297571	.0750127	1.73	0.085	-.0182178	.2777319
NW_Region	-.4707799	.1501748	-3.13	0.002	-.7670241	-.1745357
YK_Region	-.326935	.1641042	-1.99	0.048	-.6506573	-.0032127
EM_Region	-.3919707	.1743356	-2.25	0.026	-.7358759	-.0480654
WM_Region	-.415652	.1631342	-2.55	0.012	-.7374608	-.0938431
ET_Region	-.2911756	.1586935	-1.83	0.068	-.6042244	.0218732
LN_Region	-.3435547	.1650863	-2.08	0.039	-.6692143	-.017895
SE_Region	-.4454474	.1489222	-2.99	0.003	-.7392207	-.1516742
SW_Region	-.5729089	.1605562	-3.57	0.000	-.8896322	-.2561856
WL_Region	-.5552644	.1716026	-3.24	0.001	-.8937784	-.2167504
SC_Region	-.3658268	.1933083	-1.89	0.060	-.7471588	.0155052
NI_Region	-.4626371	.2042186	-2.27	0.025	-.8654916	-.0597827
_cons	.5466717	.1538675	3.55	0.000	.2431432	.8502003

The LPM Results 2012: 1

- The LPM model dependent variable is :

$$y = \begin{cases} 0 & \text{dRealturn} < 0 \\ 1 & \text{dRealturn} > 1 \end{cases}$$

- 79/205 observations have positive growth in real turnover, 38.5% of total.
- Regional dummies are control variables, only the North East Region has a positive effect (in the constant) all other regions have a negative effect, reflecting the fact that there are more negatives than positives for these regions. Are firms yet to grow after recession? North East leading the way?

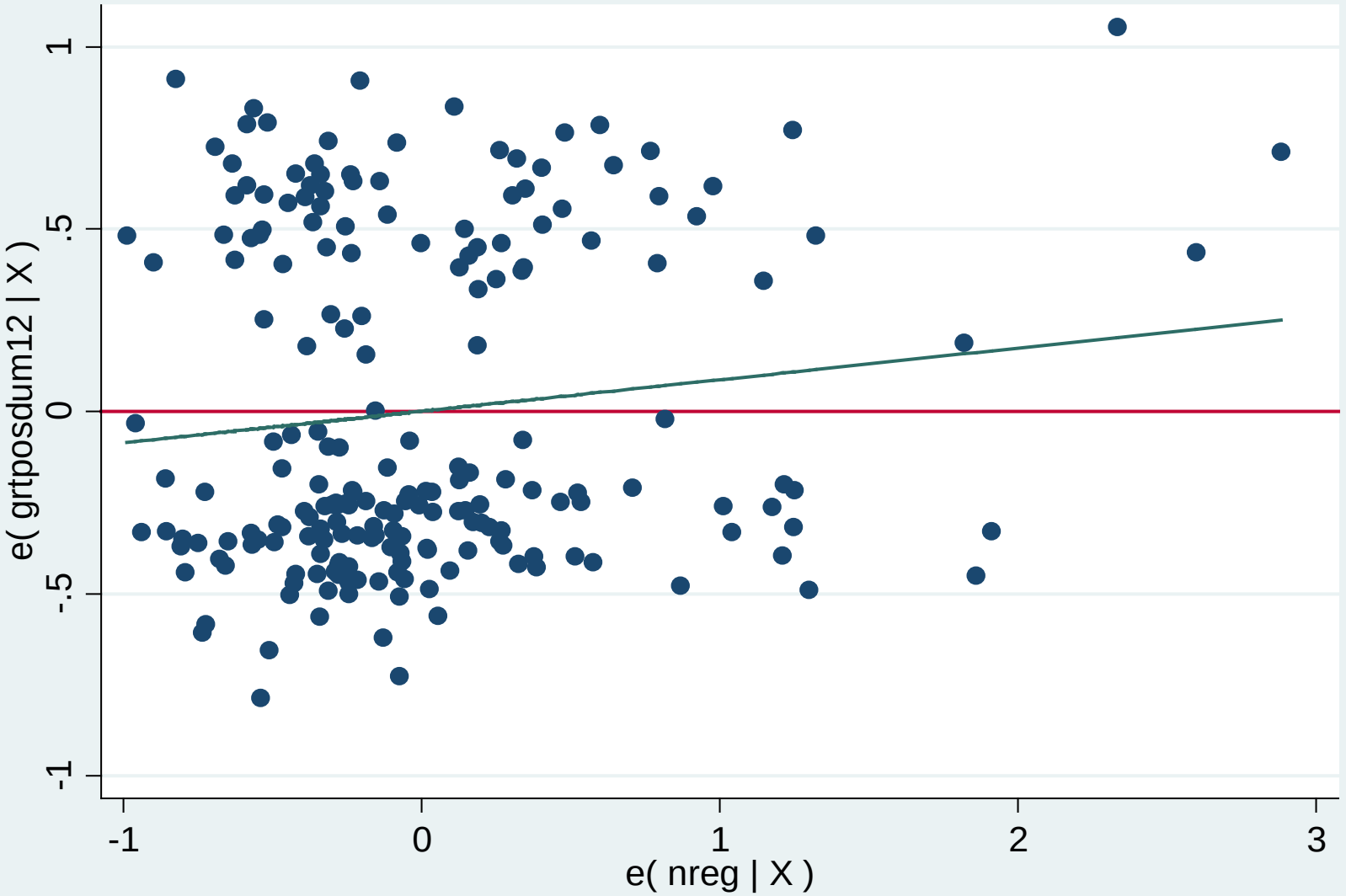
The LPM Results 2012: 2

- The estimates:

$$y = 0.55 + 0.35.\text{demp} + 0.09.\text{regulation} + 0.02.\text{man\_intensity} + 0.01.\text{rowcities}$$
$$+ 0.13.\text{SmallFirms} - 0.47.\text{NW\_Region} - .... - 0.46.\text{NI\_Region}$$

- The positive effect of employment growth on the firm, all other variables held fixed, means that an increase in demp by "1" (doubling employment) will increase the probability of firm growth by 0.35.
- The positive effect of regulation on the firm means, all other variables held fixed, that an increase in regulation by "1 unit" will increase the probability of firm growth by 0.09.

LPM variable plots (avplot)



coef = .08643758, (robust) se = .05573428, t = 1.55

LPM cross-section results for 2008

Linear regression

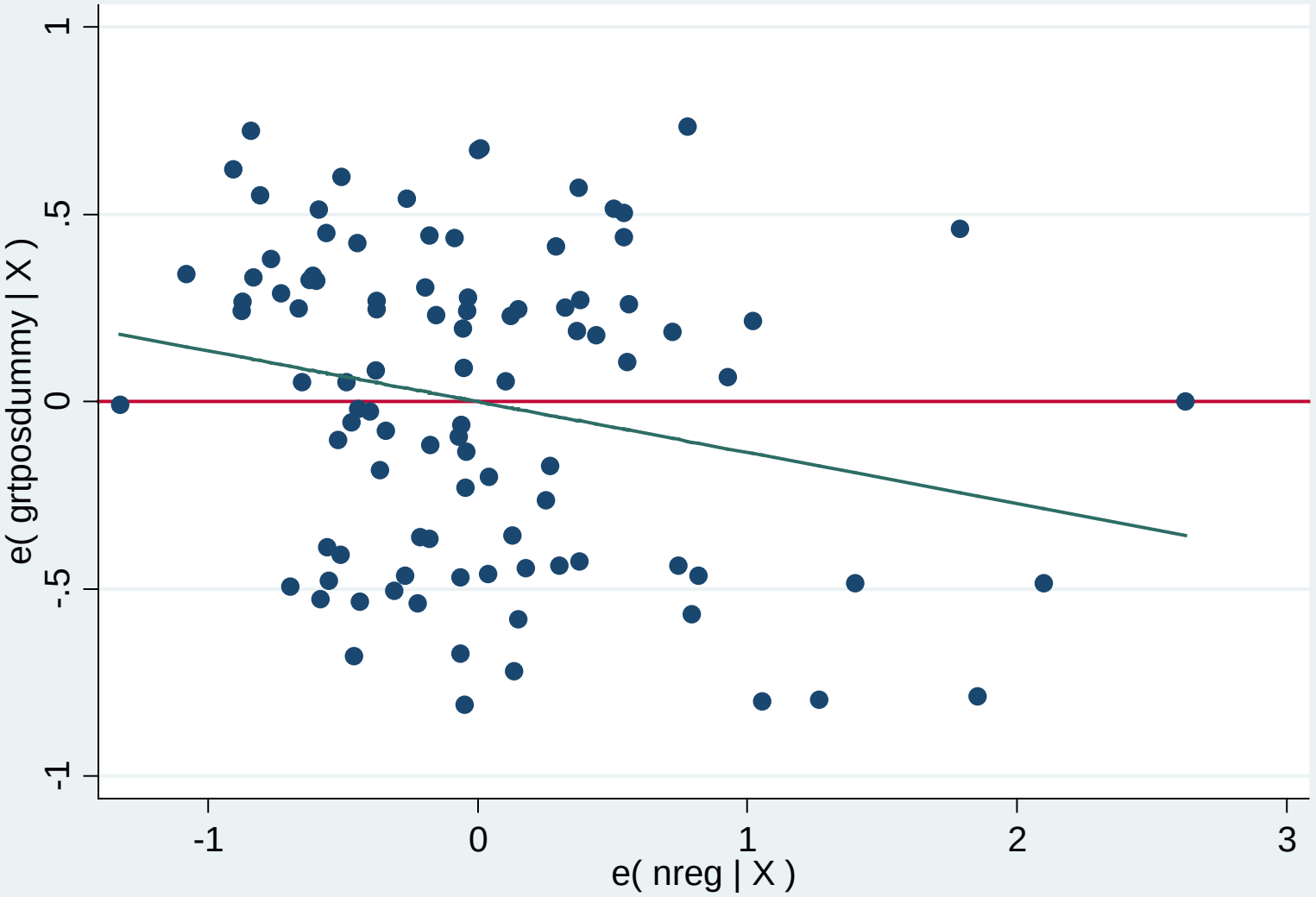
Number of obs = 98
F(15, 82) = 8.17
Prob > F = 0.0000
R-squared = 0.3051
Root MSE = .44404

grtposdummy	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
dMA4emp	.4682506	.2513024	1.86	0.066	-.03167	.9681712
FamilyOwnedD	.2986262	.1784267	1.67	0.098	-.0563214	.6535739
lag2RnDdummy	.2214935	.1107108	2.00	0.049	.0012544	.4417326
nreg	-.135813	.0679331	-2.00	0.049	-.2709536	-.0006723
npartner	.0975879	.0251067	3.89	0.000	.0476426	.1475332
NW_Region	.34344	.3146815	1.09	0.278	-.2825617	.9694417
YK_Region	.6015064	.3035222	1.98	0.051	-.002296	1.205309
EM_Region	.2878819	.3002073	0.96	0.340	-.309326	.8850899
WM_Region	.6403792	.305677	2.09	0.039	.0322903	1.248468
ET_Region	.6912016	.2884099	2.40	0.019	.1174624	1.264941
LN_Region	.3985744	.3106761	1.28	0.203	-.2194593	1.016608
SE_Region	.2849114	.304042	0.94	0.351	-.3199251	.8897479
SW_Region	.2280227	.3485319	0.65	0.515	-.4653183	.9213636
WL_Region	-.3001456	.2924279	-1.03	0.308	-.8818779	.2815868
SC_Region	.4085106	.4357921	0.94	0.351	-.4584188	1.27544
_cons	.19417	.268525	0.72	0.472	-.3400119	.7283518

The LPM Results 2008: 1

- 60/98 observations have positive growth in real turnover, 61% of total.
- Regional dummies are control variables. This time the West Midlands and Eastern Regions have a positive effect.
- The positive effect of employment growth, all other variables held fixed, means that an increase in demp by “1” (doubling employment) will increase the probability of firm growth by 0.47.
- Role here for lagged R&D activity and if the firm is Family Owned (both dummy variables).

LPM variable plots (avplot)



The LPM Results 2008: 2

- The positive effect of a firm having a partner, all other variables held fixed, means that an increase in regulation by “1 unit” will increase the probability of firm growth by 0.1.
- The negative effect of regulation on the firm, all other variables held fixed, can be interpreted as an increase in regulation by “1 unit” reducing the probability of firm growth by 0.14.
- Conclusion: when more firms are growing in 2008 the effect of more regulation is detrimental but with growth constrained conditions in 2012 regulation actually helps firm growth.

Any questions?

Panel Analysis Results

Methodological Choices: panel

- More information in a panel, firm's cross sections over time. Data collected for UK firms 2003-2012.
- Dependent variable is level of real turnover with explanatory variables: lag of real turnover, lag of employment growth, regulation web variables and year dummies for control.
- Check residuals for correlation which may indicate we have hetroscedasticity/endogenous regressors.

Panel results for linear regression


```
. xi: reg lrealturn l.lrealturn l.gremp l.ResearchDummy nreg i.year, cluster(firm)
i.year            _Iyear_2003-2012    (naturally coded; _Iyear_2003 omitted)
note: _Iyear_2004 omitted because of collinearity
note: _Iyear_2006 omitted because of collinearity

Linear regression                               Number of obs =     734
                                                F( 11,   179) =  227.46
                                                Prob > F      =  0.0000
                                                R-squared     =  0.8712
                                                Root MSE      =  .53088

                                (Std. Err. adjusted for 180 clusters in firm)
------------------------------------------------------------------------------
             |               Robust
   lrealturn |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   lrealturn |
         L1. |   .9236485   .0268042    34.46   0.000     .8707556    .9765414
             |
       gremp |
         L1. |   .1304808    .090332     1.44   0.150    -.0477718    .3087334
             |
ResearchDummy|
         L1. |   .0138671   .0452752     0.31   0.760    -.0754747    .1032088
             |
        nreg |   .0068106   .0212625     0.32   0.749    -.0351468     .048768
 _Iyear_2004 |           0  (omitted)
 _Iyear_2005 |  -.1110577   .0722317    -1.54   0.126    -.2535929    .0314775
 _Iyear_2006 |           0  (omitted)
 _Iyear_2007 |  -.0078057   .0820044    -0.10   0.924    -.1696254     .154014
 _Iyear_2008 |  -.0543979   .0501643    -1.08   0.280    -.1533874    .0445917
 _Iyear_2009 |  -.2199717   .0841959    -2.61   0.010    -.3861159   -.0538274
 _Iyear_2010 |  -.0970814   .0565052    -1.72   0.088    -.2085835    .0144207
 _Iyear_2011 |  -.0827959   .0660497    -1.25   0.212    -.2131321    .0475402
 _Iyear_2012 |  -.1825587   .0506981    -3.60   0.000    -.2826016   -.0825158
       _cons |   .8131048   .2547625     3.19   0.002     .3103806    1.315829
------------------------------------------------------------------------------
```

Panel linear model, check residuals


```
. pwcorr uols l.uols l2.uols l3.uols l4.uols, obs sig
```

	uols	L.uols	L2.uols	L3.uols	L4.uols
uols	1.0000				
	734				
L.uols	-0.1142	1.0000			
	0.0096				
	513	580			
L2.uols	-0.0642	-0.1756	1.0000		
	0.2118	0.0005			
	380	392	452		
L3.uols	0.0302	-0.1035	-0.1803	1.0000	
	0.6156	0.0805	0.0017		
	279	286	300	347	
L4.uols	0.0797	0.0554	-0.1521	0.0558	1.0000
	0.2524	0.4322	0.0264	0.4052	
	208	203	213	225	258

Some correlation over time indicating there is unobserved individual heterogeneity.

Panel results for random effects


```
Random-effects GLS regression                   Number of obs      =       734
Group variable: firm                            Number of groups   =       180

R-sq:  within  = 0.2224                         Obs per group: min =         1
       between = 0.9467                                        avg =       4.1
       overall = 0.8711                                        max =         8

                                                Wald chi2(11)      =   2806.22
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.0000
```

lrealturn	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lrealturn						
L1.	.8879361	.0169193	52.48	0.000	.8547749	.9210973
gremp						
L1.	.1070149	.0517502	2.07	0.039	.0055865	.2084434
ResearchDummy						
L1.	.0096966	.0540483	0.18	0.858	-.096236	.1156292
nreg	.0146734	.0283853	0.52	0.605	-.0409608	.0703076
year						
2006	.1074132	.1029958	1.04	0.297	-.0944549	.3092812
2007	.107624	.0971568	1.11	0.268	-.0827998	.2980478
2008	.0560725	.0943269	0.59	0.552	-.1288048	.2409498
2009	-.1152881	.094325	-1.22	0.222	-.3001617	.0695855
2010	-.0126933	.0927456	-0.14	0.891	-.1944713	.1690846
2011	.0122875	.0905005	0.14	0.892	-.1650903	.1896653
2012	-.0869263	.0892042	-0.97	0.330	-.2617633	.0879108
_cons	1.029238	.1756874	5.86	0.000	.6848973	1.373579

sigma_u	.23244199	
sigma_e	.47060341	
rho	.19611577	(fraction of variance due to u_i)

Panel results for fixed effects

```
Fixed-effects (within) regression               Number of obs      =       734
Group variable: firm                            Number of groups   =       180

R-sq:  within  = 0.2389                         Obs per group: min =         1
       between = 0.9400                                        avg =       4.1
       overall = 0.8598                                        max =         8

                                                F(11,543)          =     15.50
corr(u_i, Xb)  = 0.8806                         Prob > F           =    0.0000
```

lrealturn	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lrealturn						
L1.	.4515655	.0405512	11.14	0.000	.3719091	.5312219
gremp						
L1.	.19631	.0542408	3.62	0.000	.0897625	.3028575
ResearchDummy						
L1.	.0983449	.0703026	1.40	0.162	-.0397535	.2364433
nreg	.0277896	.0329684	0.84	0.400	-.0369716	.0925509
year						
2006	.0815275	.0984298	0.83	0.408	-.1118224	.2748773
2007	.1257978	.0932484	1.35	0.178	-.0573739	.3089695
2008	.106874	.0914909	1.17	0.243	-.0728455	.2865935
2009	-.0177054	.0919414	-0.19	0.847	-.1983098	.1628989
2010	.0005171	.0908741	0.01	0.995	-.1779908	.1790249
2011	.0384351	.0894718	0.43	0.668	-.1373181	.2141883
2012	-.0502105	.0898978	-0.56	0.577	-.2268005	.1263795
_cons	4.973846	.3738668	13.30	0.000	4.239444	5.708248
sigma_u	.95730763					
sigma_e	.47060341					
rho	.80537261	(fraction of variance due to u_i)				

Panel result: Hausman Test

```
. hausman . random
```

	Coefficients (b)	(B)	(b-B)	sqrt(diag(V_b-V_B))
	.	random	Difference	S.E.
L.lrealturn	.4515655	.8879361	-.4363706	.0368529
L.gremp	.19631	.1070149	.0892951	.0162476
L.Research~y	.0983449	.0096966	.0886483	.0449582
nreg	.0277896	.0146734	.0131163	.0167687
2006bn.year	.0815275	.1074132	-.0258857	.
2007.year	.1257978	.107624	.0181738	.
2008.year	.106874	.0560725	.0508015	.
2009.year	-.0177054	-.1152881	.0975827	.
2010.year	.0005171	-.0126933	.0132104	.
2011.year	.0384351	.0122875	.0261476	.
2012.year	-.0502105	-.0869263	.0367158	.0111453

```
                b = consistent under Ho and Ha; obtained from xtreg
     B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                 chi2(11) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =      150.26
                Prob>chi2 =      0.0000
                (V_b-V_B is not positive definite)
```

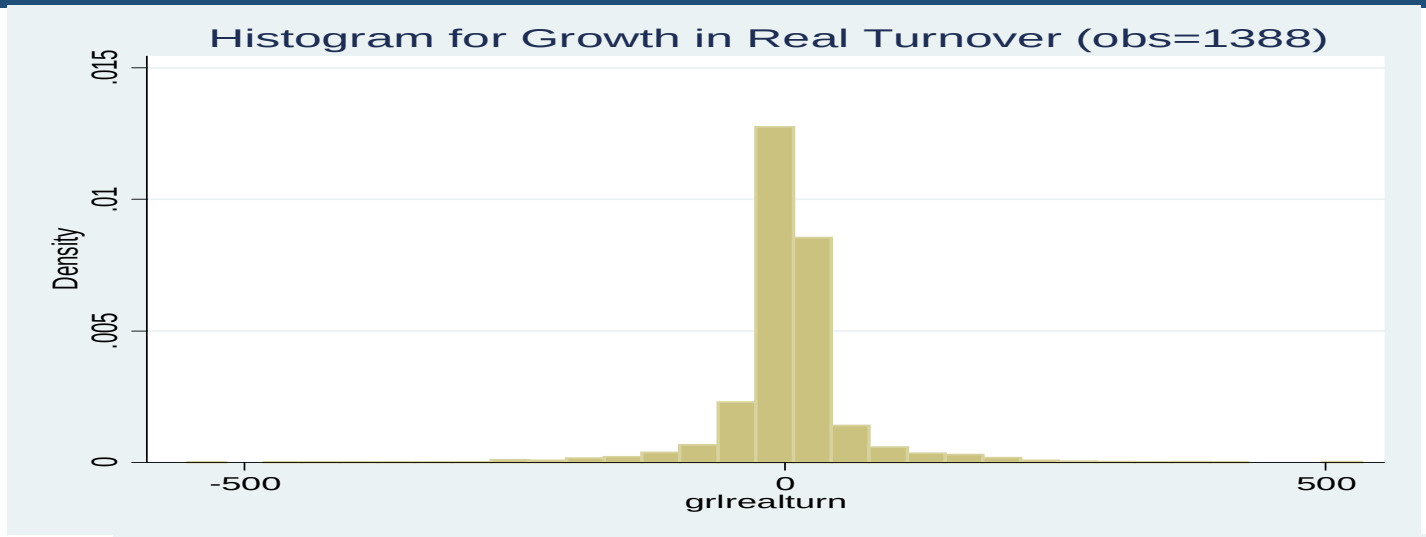
The Hausman test rejects the null that the unobserved heterogeneity is uncorrelated with the regressors, fixed effects estimates (b) are significantly different from random effects estimates (B). The fixed effects specification is preferred.

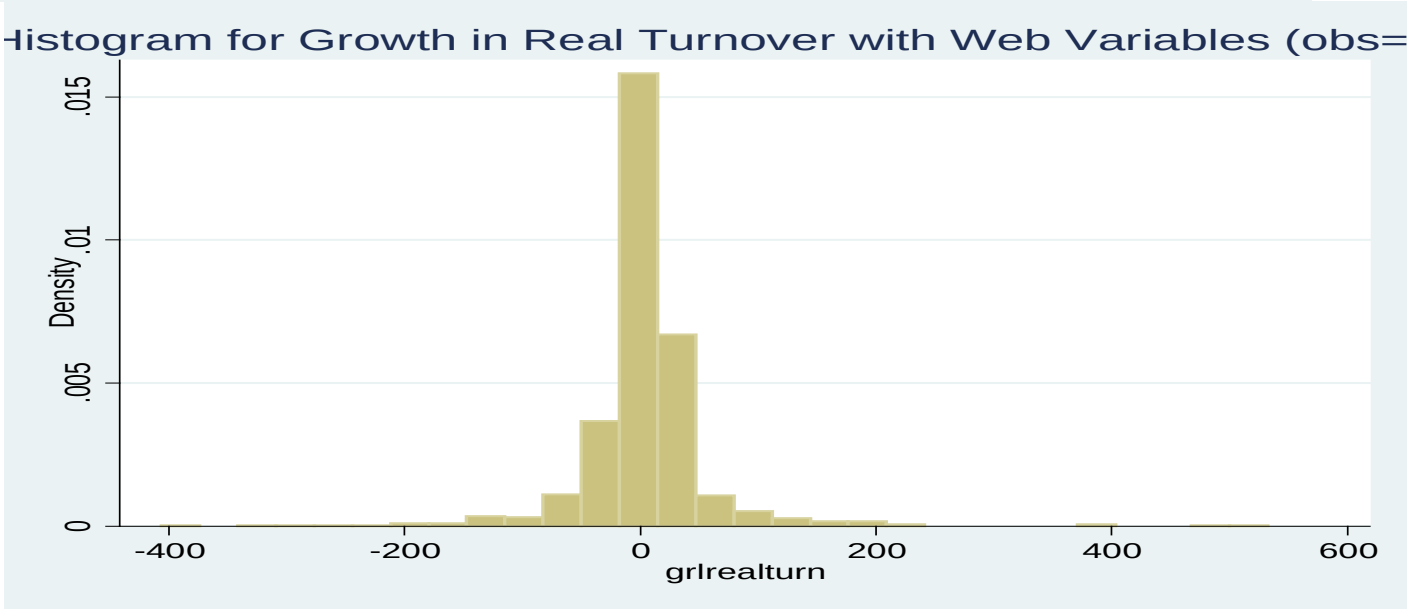
Panel results for fixed effects, robust


```
Fixed-effects (within) regression               Number of obs      =       734
Group variable: firm                            Number of groups   =       180

R-sq:  within  = 0.2389                         Obs per group: min =         1
       between = 0.9400                                        avg =       4.1
       overall = 0.8598                                        max =         8

                                                F(11,179)          =      4.87
corr(u_i, Xb)  = 0.8806                         Prob > F           =    0.0000

                                    (Std. Err. adjusted for 180 clusters in firm)
```

lrealturn	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
lrealturn						
L1.	.4515655	.1145017	3.94	0.000	.2256186	.6775124
gremp						
L1.	.19631	.0879036	2.23	0.027	.0228494	.3697706
ResearchDummy						
L1.	.0983449	.0649917	1.51	0.132	-.0299035	.2265933
nreg	.0277896	.0272805	1.02	0.310	-.0260431	.0816223
year						
2006	.0815275	.0583166	1.40	0.164	-.033549	.196604
2007	.1257978	.0644607	1.95	0.053	-.0014028	.2529984
2008	.106874	.0596357	1.79	0.075	-.0108054	.2245534
2009	-.0177054	.0863209	-0.21	0.838	-.188043	.1526321
2010	.0005171	.0841387	0.01	0.995	-.1655143	.1665484
2011	.0384351	.0915078	0.42	0.675	-.1421378	.2190079
2012	-.0502105	.098302	-0.51	0.610	-.2441904	.1437694
_cons	4.973846	1.041117	4.78	0.000	2.919405	7.028287
sigma_u	.95730763					
sigma_e	.47060341					
rho	.80537261	(fraction of variance due to u_i)				

Descriptive Statistics





xtLogit Panel: robust standard errors


```
GEE population-averaged model                   Number of obs      =       685
Group variable:                        firm     Number of groups   =       169
Link:                                 logit     Obs per group: min =         1
Family:                            binomial                    avg =       4.1
Correlation:                   exchangeable                    max =         8
                                                Wald chi2(11)      =     35.28
Scale parameter:                          1     Prob > chi2        =    0.0002

                                  (Std. Err. adjusted for clustering on firm)
```

grtposdummy	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
grtposdummy						
L1.	.3074783	.1741311	1.77	0.077	-.0338123	.648769
gremp						
L1.	.2340971	.34145	0.69	0.493	-.4351326	.9033267
ResearchDummy						
L1.	.1941359	.1737334	1.12	0.264	-.1463752	.5346471
nreg	.0888353	.1088576	0.82	0.414	-.1245217	.3021924
year						
2006	.5426736	.4484369	1.21	0.226	-.3362466	1.421594
2007	-.0244769	.3729734	-0.07	0.948	-.7554913	.7065375
2008	.1197624	.3989721	0.30	0.764	-.6622087	.9017334
2009	-.8725068	.4145433	-2.10	0.035	-1.684997	-.0600169
2010	.0759048	.3777152	0.20	0.841	-.6644033	.8162129
2011	.0084463	.3750964	0.02	0.982	-.7267292	.7436218
2012	-.7999333	.3567877	-2.24	0.025	-1.499224	-.1006423
_cons	.0691575	.3456386	0.20	0.841	-.6082817	.7465967

Panel for GGS: problems

- The UK firm dataset only captures a small number of firms. These firms have either extreme or small growth rates. Quantile regression not suitable as in Coad and Rao (2011).
- Limited dependent variable model works well for cross-sections but is less reliable in a panel.
- The amount of individual firm hetrogeneity in our sample requires a fixed effects models.