To what extent can the innovation drivers of regional competitiveness explain the GDP difference between the North and South of Italy?

Arnau Bartra
Santi Falco
Andrea Rachkov
Anouk Talen

Utrecht University

December 19, 2017

I. Introduction

Competitiveness is a multi-level concept widely discussed by academics and politicians. Its importance partly derives from the fact that competitiveness is seen as a means "for achieving and sustaining economic growth, contended living standard and well-being of people" (Borozan, 2008). However, due to the difficulty of clearly defining it, competitiveness has given rise to some conceptual issues. As the focus of this paper is regional competitiveness, the synthesis of definitions and theory will be discussed based on the regional level.

Some definitions of regional competitiveness emphasize the idea of output-related factors (e.g. productivity), which are relevant on the firm level. Others focus on the concept of prosperity, which is relevant for the residents of regions (Meyer-Stamer 2008, Bristow 2005). As both elements are important, Dijkstra et al. (2011) manage to combine the firm- and resident-perspective of regional competitiveness by defining it as:

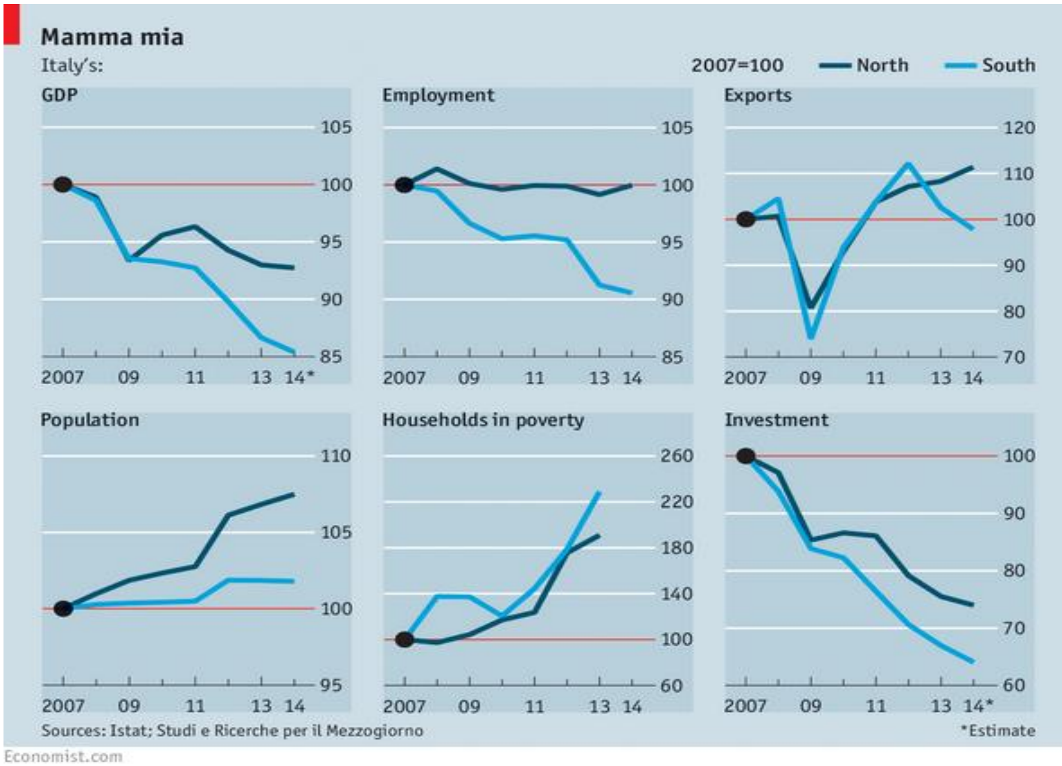
> *"…the ability to offer an attractive and sustainable environment for firms and residents to live and work".*

This definition reveals the importance of the concept as it alludes to the potential policy implications of regional competitiveness. By understanding regional competitiveness and its drivers, policy-makers could more effectively harness, foster, and sustain such environments.

In examining the EU Regional Competitiveness Index (RCI) Report, it is apparent that there may be strong differences among regions within the same country. Italy is an example of such a case. One can acknowledge that these differences do in fact exist, particularly between the northern and southern regions. The former are mostly in stage 4 or 5 (out of 5) of development and the latter in stage 2 or 3 (out of 5) according to the RCI report.

Since the Italian unification in 1861, there were pronounced differences between the North and South Italy, due to different histories and cultures. Traditionally, the North has been more industrialized than the South where the primary sector, especially agriculture, was the leading sector. The lack of infrastructure in the south of Italy shifted the trade center to the north, enabling greater economic growth in those regions. Furthermore, geographic location has had an influence where northern regions are closer to Continental Europe, thus fostering greater trade and broader knowledge exchange with foreign countries (Smith, Dennis M. 1997).

Nowadays it is possible to explain the difference between the North and South Italy with the use of different competitiveness drivers. The charts below show how various rates of crucial factors changed between 2007 to 2014, distinguishing between northern and southern regions.



Source: The Economist

As shown in the above charts, it is clear that northern regions are more competitive than the southern regions. There is a noticeable gap between the different rates, particularly in the latter years with the exception of the period during the subprime crisis.

Since 2007, the drop in the GDP level from the southern region declined at a rate double of northern regions. Although the employment rate is relatively steady in the north of the country, it collapsed in the south. In the period 2007-2014, roughly 943,000 Italians lost their job in which 70% of them were from south. The population growth rate supports our analysis. Due to the favoured conditions in northern regions, especially seen in the labour market, there had been a large migration from the South towards the North or abroad. However, this has resulted in a lack of human resources in South Italy. Despite evidence of diminishing investments in Italy, the north has continuously invested more than the south (The Economist, 2015). A plausible reason for this difference may be a result of the innovation drivers of regional competitiveness.

This group of drivers is crucial for the most advanced and competitive regions where technology and research play a predominant role in the regional growth and development. Thus, high investments in R&D, a well-functioning technological infrastructure, a high rate of advanced technologies in business operations or other factors in the technology and innovation field may have an influence on the competitiveness of a region. Since the degree of competitiveness is related to productivity, the presence of these factors in various levels can justify an inhomogeneous GDP level among different regions of the same country.

Given the previous considerations, the aim of this paper is to investigate to what extent the innovation drivers of regional competitiveness can explain the GDP difference between the North and South of Italy.

II. Methods

So far, the northern Italian regions have proven to be more competitive than the southern regions. Whilst there are various factors that could potentially explain this occurrence, our main focus will be based on the factors that belong to the innovation group.

In order to achieve our goal, we divided the twenty-one Italian regions in two groups: northern regions (Piemonte, Valle D'Aosta, Liguria, Lombardia, Provincia Autonoma di Trento, Provincia autonoma di Bolzano, Veneto, Friuli-Venezia Giulia, Emilia-Romagna,Toscana, Umbria, Marche) and southern regions (Lazio, Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria, Sicilia, Sardegna).

Using the data of the two groups, we formed two different regressions to test to what extent innovation drivers have an influence on the GDP per capita in the northern and southern regions of Italy. Innovation drivers will only have effect if the regions are sufficiently developed. Therefore, we expect that the innovation drivers will have a positive effect on the northern regions, and that the innovations drivers have less or no effect on the southern regions, since the southern regions are not developed enough yet.

The following variable was chosen to be the dependent variable in the regressions. The rationale behind choosing it is discussed below.

GDP per capita: GDP per capita is simply a measure of GDP per person in a country (i.e. GDP/Population). It was chosen for two reasons. Firstly, it reflects the economic development of a country. Secondly, although it is not the best measure of well-being, GDP per capita is still a viable indicator of standard of living in a country (United Nations, 2000).

The following innovation drivers were used as the independent variables in the regressions. The rationale behind choosing them are discussed below.

GVA in the K-N sectors: Gross value added (GVA) is the "value of output minus intermediate consumption". More importantly "it measures the contribution to GDP", in this case from the the K-N sectors (i.e. financial, real-estate, professional, scientific, and support activities sectors) (OECD, 2011). This particular GVA is calculated as percentage of total GVA.

High-tech patents and ICT patents: Patents formed by research, human creativities and inventions are an instrument to ensure firms maintain competitive advantage. Research has shown that "there exist positive relationship in long run between quarterly growth of patents and quarterly GDP growth" (Josheski and Koteski, 2011). Particularly in ICT and high-tech fields, new products developed from human knowledge and creativities enable greater efficiency and productivity of different industries, leading to higher GDP (Phirouzabadi, 2013). Both types of patents are measured as number of patent applications per million inhabitants.

Core creative class employment: According to Richard Florida (2002), the creative class comprises of workers whose economic function is to develop ideas, technology or creative content. It also includes knowledge intensive workers whose job is to create new solutions to problems. Core creative class employment is believed to provide economic growth to countries where benefits include innovation, new technologies and regional growth (Glaeser, 2005). Core creative class employment is measured as a percentage of the population aged 15 through 64.

Broadband: Household access to broadband increases real income, which directly leads to GDP growth (Katz, 2012). This variable is measured the percentage of total households with access to broadband.

The data that was taken from the OECD and Eurostat. We used part of the same data that was used to make the EU Regional Competitiveness Index RCI in 2013.

III. Results

The models were constructed using OLS estimators. To obtain the best linear unbiased estimators of our coefficients, we tested the necessary assumptions, which can be found in the appendix. These next few sections convey the outcome of our regressions and seek to explain the expected and unexpected results.

Preliminary Run of Regressions

The first run of the regressions for the North and the South included the log of GDP per capita regressed against households with access to broadband, gross value added in the K-N sectors, core creative class employment, tech patents, ICT patents, intramural R&D expenditure, the aggregated basic sub-index, and the aggregated efficiency sub-index. The rationale behind using the aggregate sub-indices was to control for basic and efficiency drivers. However, due to collinearity these variables were taken out of the regression. Furthermore, intramural R&D expenditure was also taken out of the regression due to the fact that it is measured as percentage of GDP.

GVA is also measured as part of the GDP, but we decided to keep it in the model because GVA is jointly significant with all the other independent variables in the model of the North. It is not significant in the model in the South, but we have reason to believe that this is because of the small sample size since the p-values are not extremely insignificant. The p-values in the South are all around the 0.10 (see appendix Figures 3 and 4). Another reason that we decided to keep GVA in our model is that we are not talking about the GDP of all the sectors, but specifically the GDP of sectors which are important to innovation as percentage of the total GDP of the region. Therefore, GVA gives us a way to measure how important the development of these sectors is in the North and South of Italy.

Final Run of Regressions

The final regressions include the log of GDP per capita regressed against households with access to broadband, gross value added in the K-N sectors, core creative class employment, tech patents, and ICT patents (see appendix Figures 1 and 2).

$$\textit{North}:$$
$$lgdpcap = \beta_0^N + \beta_1^N broadband + \beta_2^N gva + \beta_3^N creative + \beta_4^N techpat + \beta_5^N ictpat$$

$$\textit{South}:$$
$$lgdpcap = \beta_0^S + \beta_1^S broadband + \beta_2^S gva + \beta_3^S creative + \beta_4^S techpat + \beta_5^S ictpat$$

North	South
$\beta_0^N = -5.753$	$\beta_0^S = -15.7$
$\beta_1^N = 0.172$	$\beta_1^S = -0.027$
$\beta_2^N = 0.357$	$\beta_2^S = 0.962$
$\beta_3^N = -0.721$	$\beta_3^S = 1.371$
$\beta_4^N = -1.378$	$\beta_4^S = 2.056$
$\beta_5^N = 0.116$	$\beta_5^S = -2.482$

Significance Testing

Since the regressions contain multiple population parameters, the F-test with 5% significance level was used to determine the significance of the results.

The following null alternative hypotheses were tested for the North:

$$H_0: \beta_1^N = \beta_2^N = \beta_3^N = \beta_4^N = \beta_5^N = 0$$
$$H_A: H_0 \text{ not true}$$

This resulted in a p-value of 0.1005 as can be seen in Figure 5 of the appendix. Since 0.1005>0.05 H_0 was not rejected, which implies that the results are insignificant.

The following null alternative hypotheses were tested for the South:

$$H_0: \beta_1^S = \beta_2^S = \beta_3^S = \beta_4^S = \beta_5^S = 0$$
$$H_A: H_0 \text{ not true}$$

This resulted in a p-value of 0.215 as can be seen in Figure 6 of the appendix. Since 0.215>0.05, H_0 was not rejected, which implies that the results are insignificant. However, due to the prior economic reasoning behind each one of the variables, we believe to have enough theoretical ground to motivate our subsequent interpretation of the results.

Basic Interpretation of the Independent Variables

This section focuses on the basic interpretation of the the independent variables. Note that for every interpretation the rest of the independent variables are constant (i.e. ceteris paribus assumption).

GVA seems to be the most important contributor for the GDP per capita in the North. When GVA goes up by one percent, the GDP per capita of the North goes up by 35.7% and the GDP per capita of the South goes up by 96.2%.

While GVA is a big contributor to the GDP per capita of the southern regions, there are two other independent variables that have an even bigger influence on the GDP per capita in the South: core creative class employment and high-tech patents. When the percentage of core class creative employment in the South increases by one percent, the GDP per capita increases by 137%; while in the North a one-percent increase leads to a decrease in GDP per capita by 72.1%. High-tech patents have an even bigger effect on the GDP per capita in the South. One patent more per million inhabitants leads to an increase of

205.6% of GDP per capita, while in the North an additional patent per million inhabitants has a negative effect on GDP per capita of 137.8%

While high-tech patents have a large positive effect in the South, ICT-patents have a large negative effect on the GDP per capita. When patent applications per million inhabitants increases by one, then the GDP per capita goes down by 248.2%. On the other hand, ICT-patents do have a positive effect in the North. There, an increase of one patent per millions inhabitants increases the GDP per capita by 11.6%.

Household access to internet also has a negative effect on the GDP per capita in the South. When the household access to internet goes up by one percent, the GDP per capita goes down by 2.7 %. The access to internet does have a positive effect on the GDP per capita of the northern regions, as it goes up by 17.2% per unit increase.

As one can see, the majority of the values of the coefficients of the independent variables in the regression for the southern regions are greater than the values of the coefficients of the independent variables of the regression for the northern regions (with the exception of broadband and ICT-patents). This means that an increase in the independent variable in the South has a greater effect of GDP per capita than an increase in an independent variable has in the North. This is the opposite of what we had expected. A possible reason for this will be discussed in the next section.

Another observation is that independent variables that have a positive effect in the North, have a negative effect in the South, and vice versa (with the exception of GVA).

Explanation of the Results

This section is dedicated to providing a possible explanation for the regression results.
We start with the household access to internet. The household access to broadband has a positive effect on the GDP per capita in the North, while it has a small negative effect on GDP per capita in the South. The positive effect in the North is logical as access to internet allows people to share and transfer data at high speeds, make faster transactions and other activities that have a positive impact on economic growth (Manyika and Roxburgh, 2011) . Although the negative coefficient for household access to broadband in the South seemed inexplicable at first, further research yielded a possible justification. ICT specialist Victor Mulas (2011) hypothesizes that "not all economies are equally prepared to absorb broadband and embrace it to reap potential benefits". He refers to absorptive capacity and the fact that it may vary per economy causing different impacts from the access to broadband (Mulas, 2011).

According the model, ICT-patents have a positive effect on the GDP per capita in the North and a negative effect on the GDP per capita in the South. ICT services help firms to communicate and connect with the outer world as well as the inner world of a company. New ideas concerning ICT can help firms work more efficiently and to better process data. A good reason why the ICT-patent applications have a positive effect in the northern regions is that the benefits of the patents weigh out the costs of patents. In a developed region the firms are in a good business environment, which enables firms to grow. When one grows as a firm, one is dealing with more data and the business processes becomes more elaborate. New ICT innovations are then very useful since they can lead to more efficiency. For the southern regions this is not the case. The costs of the ICT investment is higher than the benefits. In the less developed South, the business environment is not optimal and firms are not necessarily growing. When a firm is not growing, it does not need the newest innovations with respect to ICT. A firm could even be better off implementing older versions of ICT technologies rather than implementing the newest ones. So, for a firm to use its money for ICT-patents can have the opposite effect. It costs substantial amount of money to invest in it, but these innovations do not necessarily make every company more efficient.

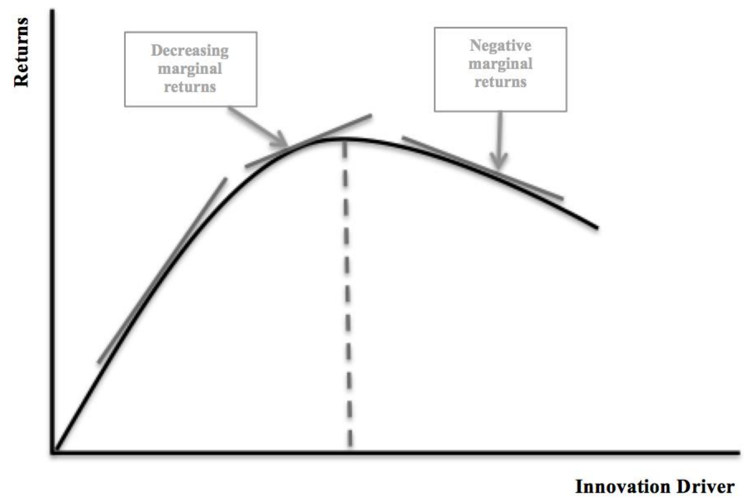
Another explanation why the ICT patents have a positive effect in the North and a negative effect in the South is that the sectors that rely heavily on ICT are more located in the North, and that in sectors most present in the South are sectors in which ICT is not very important, or maybe not even important at all. This way ICT patent applications have a positive effect on the GDP in the North, while in the South the ICT patent applications are a cost to the company instead of a benefit, and therefore reduce GDP.

The difference in impact of GVA between the North and the South of Italy could be explained by the effect of decreasing marginal returns. The marginal returns imply that the greater the increase in GVA, the less one additional unit of the GVA will add to GDP per capita. This could be due to the fact that the North is more developed than the South, so an augmentation of innovation drivers in the South could have a higher marginal return than in the North. This is exactly what our results show. While a one percent increase of the GVA in the North leads to an increase of 35.7% in GDP per capita, one percent increase of the GVA in the South leads to an increase of 96.2%.

For the tech-patents and the core creative class employment, the value in the North is negative and the value in the South is positive. This could be explained by negative marginal returns. Thus, at some point, an increase in a certain innovation driver could be so much that not only do the marginal returns decrease, but they also turn negative. Since regions in the North are already in the more advanced stages of development, too big of an increase of the two drivers can lead to a negative effect on the GDP per capita. In the South, the increase leads to a positive effect on GDP per capita.

For tech-patents this effect could be logically explained. Too little patents can mean not enough innovation. If companies do not innovate, this could have negative effect on regional economies, as innovation is an important contributor to productivity and thus GDP per capita (Hall, 2011). So, more patents can increase “productivity and economic growth” (Crosby, 2002). On the other hand, too many patents could stiffen competition, since other companies can not profit from the innovation. This leads to the creation of monopolies and can result in less innovation and less companies, which ultimately has a negative effect on GDP per capita.

For the creative class employment a similar rationale can be used. If more people between the ages of 14 and 65 are part of this section of the workforce, there is a greater amount of ‘creative’ input. This has a positive influence on GDP per capita. However, if there are too many people in the core creative class employment, there may be too little workers in other important sections of the workforce. Thus, this could have a negative effect on GDP per capita.



VI. Conclusions

We have shown that the innovation drivers do indeed have a different impact in the North and South of Italy. However, the effect is different than what we had initially expected (i.e. we expected the impact of the variables to be greater in the northern regions). While the northern regions are performing better when households have access to broadband and have more ICT-patent applications, the southern regions have a greater impact on their GDP per capita from higher GVA, increased core creative class employment, and more tech-patents. The difference in the effect of these innovation drivers between the northern and southern regions can be attributed to the differing levels of development in these regions.

As a result of these differing levels of development certain variables assist the competitiveness of the northern regions against competitive in Europe, whilst other variables build the competitiveness of the Southern regions in their capacity to increase the development of their infrastructure. For example, the Northern regions can benefit greater from access broadband, as they already have an advanced infrastructure that allows them to absorb these improvements to help increase the level of competitiveness. Meanwhile, in the south variables such as increased core creative class employment assist in increasing per capita GDP, as these are necessary to capitalize on the benefits of innovation drivers.

The model shows that some of the innovation drivers have a low marginal return in the North of Italy, while in the South of Italy the marginal returns are higher. Much in the same way that adding more workers to a machine than the machine can support will not increase output, or may even decrease output, the marginal return in the north is lower since an increase in the already very advanced variables does not result in a large increase in GDP per capita. On the other hand, since it is a less developed area, the south benefits from large marginal returns as it continues to work towards the levels of development of the North.

Furthermore, the model shows that for the drivers: technological patents and core creative class employment there exist an optimal level. Too low values are not good in this case, but too high values are not good for economic growth either.

Limitations

Unfortunately, our model is also subject to significant limitations. Firstly, Italy is a country consisting of only 21 NUTS2 regions. This implies that there are fewer observations in the sample group than the suggested minimum of 30 observations. As our aim was to explore the difference in impact of the variables between the North and South, the sample sizes decreased even further (North had 13 observations while the South had 8 observations). As a result, only one or two variables in the model turned out to be statistically significant.

Furthermore, the inability to include the aggregated values of the basic and efficiency sub-indices into the regressions due to unexpected collinearity may have decreased the reliability of the model. Without them we could not control for their effect (especially when discussing the South).

V. References

Annoni, P. and Dijkstra, L. (2013). EU Regional Competitiveness Index RCI 2013. *JRC Scientific and Policy Reports.*

Borozan, D. (2008). Regional Competitiveness: Some Conceptual Issues and Policy Implications. *Interdisciplinary Management Research IV* (pp.50-61). Osijek: Faculty of Economics in University of Osijek.

Bristow, G. (2005). Everyone's a 'winner': problematising the discourse of regional competitiveness. *Journal of Economic Geography*, 5(3), 285-304.

Crosby, M. (2000). Patents, Innovation, and Growth. *The Economic Record*, 76(243), 255-262.

Dijkstra, L. et al. (2011). A new regional competitiveness index: Theory, Methods, and Findings. *European Union Regional Policy Working Papers.*

Florida, R. (2002). *The Rise of the Creative Class.* (1st ed.). New York: Basic Books.

Hall, B. (2011). Innovation and productivity. *Nordic Economic Policy Review.*

Josheski, D. and Koteski, C. (2011). The causal relationship between patent growth and growth of GDP with quarterly data in G7 countries: cointegration, ARDL, and error correction models. *Munich Personal RePEc archive.* Retrieved 06 December, 2016, from https://mpra.ub.uni-muenchen.de/33153/

Katz, R. (2012). The Impact of Broadband on the Economy: Research to Date and Policy Issues. *Broadband Series.*

OECD (2001). *OECD Glossary for Statistical Terms.* Retrieved 02 December, 2016, from https://stats.oecd.org/glossary/detail.asp?ID=1184

Meyer-Stamer, J. (2008). Systemic Competitiveness and Local Economic Development. *In Bodhanya, S. (Ed), Large Scale Systemic Change: Theories, Modeling, and Practices* (pp.217-240). Duisburg.

Manyika J. and Roxburgh C. (2011). The great transformer: the impact of the Internet on economic growth and prosperity. *McKinsey Global Institute.*

Mulas, V. (2011). Why does broadband not always have an impact on economic growth? Retrieved 10 December, 2016, from http://worldbank.org/ic4d/why-broadband-does-not-always-have-an-impact-on-economics -growth

Phirouzabadi, A. (2013). Examining and Analyzing the Relation between Patents with GDP and GDP per capita: Studying 33 countries as case studies. Retrieved 06 December, 2016, from https://www.researchgate.net/publication/261788716_Examining_and_ Analyzing_the_Relation_between_Patents_with_GDP_and_GDP_per_Capita Studying_33_countries_as_case_studies

Smith, Dennis M. (1997). *Modern Italy: A Political History*.

Glaeser, E. (2005). Review of Richard Florida's "The rise of the creative class". *Regional Science and Urban Economics*, 35(5), 593-596.

United Nations (2000). Indicators of Sustainable Development: Guidelines and Methodology. http://www.un.org/esa/sustdev/publications/indisd-mg2001.pdf

VI. Appendix

i. Figure 1: Regression (South)

```
. reg lgdpcap techpat ictpat broadband gva creative

      Source |       SS           df       MS      Number of obs   =         8
-------------+----------------------------------   F(5, 2)         =      3.95
       Model |  6.12616422         5  1.22523284   Prob > F        =    0.2145
    Residual |  .621128001         2     .310564   R-squared       =    0.9079
-------------+----------------------------------   Adj R-squared   =    0.6778
       Total |  6.74729222         7  .963898888   Root MSE        =    .55728

------------------------------------------------------------------------------
     lgdpcap |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     techpat |   2.055477   .6489023     3.17   0.087    -.7365247    4.847478
      ictpat |  -2.481759    .742782    -3.34   0.079    -5.677692    .7141735
   broadband |   -.027188   .0683526    -0.40   0.729    -.3212858    .2669097
         gva |   .9616506   .2369207     4.06   0.056     -.057737    1.981038
    creative |     1.3705   .7431529     1.84   0.206    -1.827029    4.568029
       _cons |      -15.7   6.228225    -2.52   0.128    -42.49788    11.09789
------------------------------------------------------------------------------
```

ii. Figure 2: Correlation Estimated Error Term (uhat) and Independent Variables (South)

```
. corr uhat broadband gva creative techpat ictpat
(obs=8)

             |     uhat broadb~d      gva creative  techpat   ictpat
-------------+------------------------------------------------------
        uhat |   1.0000
   broadband |  -0.0000   1.0000
         gva |  -0.0000   0.0131   1.0000
    creative |  -0.0000   0.8221   0.0998   1.0000
     techpat |  -0.0000  -0.3697   0.5422  -0.1816   1.0000
      ictpat |  -0.0000  -0.1482   0.7462   0.0660   0.9193   1.0000
```

iii. Figure 3: Correlation Dependent Variables and Independent Variables (South)

```
. corr lgdpcap broadband gva creative techpat ictpat
(obs=8)

             |  lgdpcap broadb~d      gva creative  techpat   ictpat
-------------+------------------------------------------------------
     lgdpcap |   1.0000
   broadband |   0.0187   1.0000
         gva |   0.6258   0.0131   1.0000
    creative |   0.0982   0.8221   0.0998   1.0000
     techpat |   0.3435  -0.3697   0.5422  -0.1816   1.0000
      ictpat |   0.3367  -0.1482   0.7462   0.0660   0.9193   1.0000
```

iv. Figure 4: Regression (North)

```
. reg lgdpcap techpat ictpat broadband gva creative

      Source |       SS           df       MS      Number of obs   =        13
-------------+----------------------------------   F(5, 7)         =      2.88
       Model |  11.8399528         5  2.36799056   Prob > F        =    0.1005
    Residual |  5.76275419         7  .823250598   R-squared       =    0.6726
-------------+----------------------------------   Adj R-squared   =    0.4388
       Total |   17.602707        12  1.46689225   Root MSE        =    .90733

------------------------------------------------------------------------------
     lgdpcap |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     techpat |  -.1376968    .115686    -1.19   0.273    -.4112507    .1358571
      ictpat |   .1156427   .1050025     1.10   0.307    -.1326487    .3639341
   broadband |   .1719337   .1284491     1.34   0.223    -.1318001    .4756675
         gva |   .3573561   .1107444     3.23   0.015     .0954872    .6192249
    creative |  -.7213236   .6647739    -1.09   0.314    -2.293264     .850617
       _cons |   -5.75328   8.164322    -0.70   0.504    -25.05883    13.55227
------------------------------------------------------------------------------
```

v. Figure 5: Correlation Estimated Error Term (uhat) and Independent Variables (North)

```
. corr uhat broadband gva creative techpat ictpat
(obs=13)

             |     uhat broadb~d      gva creative  techpat   ictpat
-------------+------------------------------------------------------
        uhat |   1.0000
   broadband |   0.0000   1.0000
         gva |  -0.0000  -0.2092   1.0000
    creative |  -0.0000   0.2190  -0.0113   1.0000
     techpat |  -0.0000  -0.2284  -0.1123   0.4226   1.0000
      ictpat |  -0.0000  -0.3738   0.0231   0.5794   0.8027   1.0000
```

vi. Figure 6: Correlation Dependent Variables and Independent Variables (North)

```
. corr lgdpcap broadband gva creative techpat ictpat
(obs=13)

             |  lgdpcap broadb~d      gva creative  techpat   ictpat
-------------+------------------------------------------------------
     lgdpcap |   1.0000
   broadband |   0.0659   1.0000
         gva |   0.7156  -0.2092   1.0000
    creative |  -0.1543   0.2190  -0.0113   1.0000
     techpat |  -0.3371  -0.2284  -0.1123   0.4226   1.0000
      ictpat |  -0.1554  -0.3738   0.0231   0.5794   0.8027   1.0000
```

vii. Figure 7: GVA Joint Significance (South)

```
. test gva broadband

 ( 1)  gva = 0
 ( 2)  broadband = 0

       F(  2,     2) =    8.26
            Prob > F =    0.1080

. test gva creative

 ( 1)  gva = 0
 ( 2)  creative = 0

       F(  2,     2) =    8.30
            Prob > F =    0.1075

. test gva techpat

 ( 1)  gva = 0
 ( 2)  techpat = 0

       F(  2,     2) =    8.57
            Prob > F =    0.1045

. test gva ictpat

 ( 1)  gva = 0
 ( 2)  ictpat = 0

       F(  2,     2) =    8.28
            Prob > F =    0.1078
```

viii. Figure 8: GVA Joint Significance (North**)**

```
. test gva broadband

 ( 1)  gva = 0
 ( 2)  broadband = 0

       F(  2,     7) =    5.45
            Prob > F =    0.0374

. test gva ictpat

 ( 1)  gva = 0
 ( 2)  ictpat = 0

       F(  2,     7) =    5.97
            Prob > F =    0.0306

. test gva techpat

 ( 1)  gva = 0
 ( 2)  techpat = 0

       F(  2,     7) =    6.82
            Prob > F =    0.0227

. test gva creative

 ( 1)  gva = 0
 ( 2)  creative = 0

       F(  2,     7) =    5.55
            Prob > F =    0.0360
```

ix. Figure 9: F-Test for Hypothesis Testing (North)

```
. test gva broadband creative ictpat techpat

 ( 1)  gva = 0
 ( 2)  broadband = 0
 ( 3)  creative = 0
 ( 4)  ictpat = 0
 ( 5)  techpat = 0

       F(  5,     7) =    2.88
            Prob > F =    0.1005
```

X. Figure 10: F-test for hypothesis testing (South)

```
. test gva broadband creative techpat ictpat

 ( 1)  gva = 0
 ( 2)  broadband = 0
 ( 3)  creative = 0
 ( 4)  techpat = 0
 ( 5)  ictpat = 0

       F(  5,     2) =    3.95
            Prob > F =    0.2145
```

x. Testing for Unbiasedness

For the sampling distribution to be centered on the true population β_i (i.e. $E(\beta_i)=\beta_i$) the following four assumptions must hold. As the first assumption of linearity cannot be tested, it must be assumed true in order to be able to use OLS estimators for the regressions. The second assumption, which calls for the error terms to have a zero mean (i.e. $E(\varepsilon_i)=0$), holds for both regressions as they each include a constant (β_0^N= -5.735 and β_0^S=-15.7). The third assumption, which states that the independent variables must be exogenous, can be determined by checking that $Corr(X_i,\varepsilon_i)=0$ for both regressions. Figures 2 and 5 in the appendix show that this does indeed hold as none of the independent variables in either regression have any correlation with the estimated error terms. The last assumption of no perfect multicollinearity also holds as none of the independent variables share a correlation of or with each other. This is shown in Figures 3 and 6 of the appendix.

xi. Testing for Efficiency

For an unbiased estimate of $Var(\beta_i)$ the last four assumptions and an additional two assumption must hold. The fifth assumption that calls for no correlation of error terms between observations (i.e. $Corr(\varepsilon_j,\varepsilon_i)=0$). This is mostly a problem for time series data. As it cannot be determined, this assumption is taken as a given. The last assumption of no heteroskedasticity states that the variance of the error term should be constant (i.e. $Var(\varepsilon_i)=\sigma^2$).

xii. Heteroscedasticity Test North

H0: no heteroskedasticity

HA: heteroskedasticity

Significance level is 5%

```
. predict uhat, resid

. gen uhat2 = uhat*uhat

. reg uhat2 techpat ictpat creative broadband gva

      Source |       SS           df       MS      Number of obs   =        13
-------------+----------------------------------   F(5, 7)         =      1.63
       Model |  1.26804998         5  .253609995   Prob > F        =    0.2695
    Residual |  1.09217887         7  .156025553   R-squared       =    0.5373
-------------+----------------------------------   Adj R-squared   =    0.2067
       Total |  2.36022885        12  .196685738   Root MSE        =      .395

------------------------------------------------------------------------------
       uhat2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     techpat |  -.0454047   .0503631    -0.90   0.397    -.1644946    .0736852
      ictpat |   .0349288   .0457121     0.76   0.470    -.0731632    .1430208
    creative |  -.5714333   .2894049    -1.97   0.089    -1.255767    .1129005
   broadband |   .0004917   .0559194     0.01   0.993    -.1317368    .1327201
         gva |  -.0090972   .0482118    -0.19   0.856       -.1231    .1049056
       _cons |   3.993303   3.554283     1.12   0.298     -4.41124    12.39785
------------------------------------------------------------------------------

. test techpat ictpat broadband gva creative

 ( 1)  techpat = 0
 ( 2)  ictpat = 0
 ( 3)  broadband = 0
 ( 4)  gva = 0
 ( 5)  creative = 0

       F(  5,     7) =    1.63
            Prob > F =    0.2695
```

$0.2695 > 0.05$. So we don't reject the zero-hypothesis and conclude that the data is homoscedastic