SERC DISCUSSION PAPER 109

Does Self-Employment Measure Entrepreneurship? Evidence from Great Britain

Giulia Faggio (SERC)
Olmo Silva (SERC, Department of Geography and Environment, LSE and IZA)

May 2012

This work is part of the research programme of the independent UK Spatial Economics Research Centre funded by a grant from the Economic and Social Research Council (ESRC), Department for Business, Innovation & Skills (BIS) and the Welsh Assembly Government. The support of the funders is acknowledged. The views expressed are those of the authors and do not represent the views of the funders.

© G. Faggio and O. Silva, submitted 2012

Does Self-Employment Measure Entrepreneurship? Evidence from Great Britain†

Giulia Faggio* and Olmo Silva**

May 2012

* Spatial Economics Research Centre (SERC), London School of Economics
** Department of Geography and Environment and Spatial Economics Research Centre (SERC), London School of Economics; and IZA Bonn.

†This work was based on data from the Business Structure Database, Community Innovation Survey and Quarterly UK Labour Force Survey produced by the Office for National Statistics (ONS) and supplied by the Secure Data Service at the UK Data Archive. The data are Crown Copyright and reproduced with the permission of the controller of HMSO and Queen's Printer for Scotland. The use of the data in this work does not imply the endorsement of ONS or the Secure Data Service at the UK Data Archive in relation to the interpretation or analysis of the data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

Abstract
Research on entrepreneurship often uses information on self-employment to proxy for business creation and innovative behaviour. However, little evidence has been collected on the link between these measures. In this paper, we use data from the UK Labour Force Survey (LFS) combined with data from the Business Structure Database (BSD), and the Community Innovation Survey (CIS) to study the relation between self-employment, business creation and innovation. In order to do so, we aggregate individual and firm-level data at the Travel-to-Work Area (TTWA) and investigate how the incidence of self-employment correlates with the density of business start-ups and innovative firms. Our results show that in urban areas a higher incidence of self-employment positively and strongly correlates with more business creation and innovation, but this is not true for rural areas. Further analysis suggests that this urban/rural divide is related to lack of employment opportunities in rural areas, which might push some workers into self-employment as a last resort option.

JEL Classifications: L26, J21, R12, R23
Keywords: Entrepreneurship; self-employment; spatial distribution

Acknowledgements
We would like to thank Philippe Bracke, Stephanie Dales, Ian Gordon, Max Nathan, Henry Overman, Will Strange and participants at the "Urban and Regional Economics" SERC/LSE seminar for comments and suggestions. We are responsible for any errors or omissions.

1. Introduction

Since the writings of Marshall (1890) and Schumpeter (1921), entrepreneurship is believed to be a key determinant of the economic success of a country or region, and a crucial factor in shaping the spatial distribution of economic activities on the national territory. Entrepreneurs are not only responsible for the creation of new firms, but also for their technological lead and success. Additionally, entrepreneurs are identified as innovators and sources of job creation (see Acs and Audretsch, 2003; and Storey, 2006). In a nutshell, entrepreneurs are engines of economic growth, and differences in levels of entrepreneurial activities bear important implication for disparities in income across countries and regions.

Unsurprisingly, policy makers devote a lot of attention to business start-up rates and small business creation, and have set in place a number of institutions aimed at promoting entrepreneurship. In the US, the federally funded Small Business Administration (SBA) agency was created in 1953 with the aim of helping Americans to "start, build and grow businesses" through programmes that facilitate access to capital and provide education, information and technical assistance. In the UK, the Department for Business, Innovation and Skills (BIS) assists small businesses through the Enterprise and Business Directorate[1] with the aim of "enabling more people (…) to start their business" and "boosting enterprises, start-ups and small business growth (…) through improved access to finance (…), and more positive business environment", as well as programmes to support people to start and manage their business. Recently, President Barack Obama presenting the American Jobs Act (September 2011) stated that "everyone knows that small businesses are where most new jobs begin. (…) So for everyone who speaks passionately about making life easier for job creators, this plan is for you".

Despite its relevance to both economic thinking and policy making, academic research on entrepreneurship is partly impaired by fundamental issues surrounding the definition of entrepreneur and the identification of entrepreneurial individuals in available data. The vast majority of the empirical work has focused on the study of self-employment. Examples include Blanchflower and Shadforth (2007) and Evans and Leighton (1989) on trends in entrepreneurship in the UK and the US, respectively; Blanchflower and Oswald (1998), Evans and Jovanovic (1989), Holtz-Eakin *et al*. (1994a) and (1994b), Hurst and Lusardi (2004) and Michelacci and Silva (2007) on the role of credit constraints; Cagetti and De Nardi (2009) and

[1] The UK government has a long history in supporting small businesses. For instance, during 1979-1983, more than one hundred measures were introduced by the Conservative administration to assist small firms. In 1979, a major initiative called the Business Expansion Scheme was introduced (then discontinued and reintroduced as the Enterprise Investment Scheme). Since then, a number of government-sponsored schemes assisting small businesses have existed with different names and administrated by different agencies/departments over time (Storey, 1994).

Carroll *et al.* (2000) on the role of taxation; and Ardagna and Lusardi (2008), Lazear (2004) and Silva (2007) on the role of skills and individual characteristics.

According to the seminal writings by Knight (1921) and Schumpeter (1942), entrepreneurs are individuals who bring innovations to the market in a process of creative destruction, strive to grow and create jobs, and bear the risk of the uncertainty surrounding entrepreneurial success (see also Kanbur, 1979; and Khilstrom and Laffont, 1979). From this perspective, not all self-employed individuals are necessarily entrepreneurial types. Alba-Ramirez (1994) and Martinez-Granado (2002) show that some self-employment spells can be explained by lack of alternative employment opportunities, and therefore identify self-employment out-of-necessity.

A similar distinction is proposed by Baumol (2011) – and supported by research sponsored by the Kauffman Foundation – who distinguishes between 'innovative' and 'replicative' (or routine) entrepreneurs. According to the author, only the former are key to an economy's long-run success by supplying new products, ideas and processes. Conversely, replicative entrepreneurs, who manage retail units and other small businesses, predominantly respond to local demand and growing population, and are therefore symptoms of a growing economy rather than causes. Nevertheless, it should be noted that routine entrepreneurship could have substantial beneficial effects on economic efficiency by promoting division of labour, by providing required goods and services and by fostering input-output linkages.

Recently, Hurst and Pugsley (2010) and Sanandaji (2010) investigate these issues directly and cast further doubts on the mapping between self-employment and entrepreneurship. Hurst and Pugsley (2010) show that the vast majority of US small businesses do not innovate, do not want to innovate, do not significantly grow in size and do not want to expand. This suggests that most US self-employed workers are hardly entrepreneurial from the perspective of innovation and job creation. Along similar lines, Sanandaji (2010) uses cross-country data to document that the correlation between the incidence of self-employment and billionaires who became rich by setting up their own business (as listed in Forbes Magazine) is negative and significant.

In this paper, we investigate these issues by combining three sources of data, namely the UK Labour Force Survey (LFS), the Business Structure Database (BSD), and the Community Innovation Survey (CIS). To start with, we use information contained in the LFS over the period 1995 to 2009 to identify individuals who are: *(i)* self-employed; *(ii)* self-employed in managerial and professional occupations; *(iii)* self-employed owning a business or a controlling majority of the business where they work; and *(iv)* self-employed individuals who are not freelancers or subcontractors (more details in Section 2). The LFS includes a large and representative sample of individuals in the UK, and these definitions of entrepreneurs are similar to those adopted by

previous studies that analyse business creation using individual-level data and take self-employment as a proxy for entrepreneurship.

Next, we use information contained in the BSD over the period 1997 to 2008 to identify how many firms are created and destroyed every year, and compute proxies for net and gross firm creation. The BSD is an administrative dataset that covers almost all business organizations in the UK, including both single and multi-plant enterprises. Finally, we use data from the CIS in 2001, 2005, 2007 and 2009 to identify firms that innovate by creating either new products or new processes of production. The CIS is part of a broader European Community data collection effort and its UK component has a sample that was chosen to be representative of small, medium and large businesses, across all UK regions and core sectors.

In order to compare the incidence of self-employment with the intensity of business start-ups and innovative behaviour, we aggregate individual-level data and firm-level data at the Travel-to-Work Area (TTWA) level. These areas are functional geographical units constructed by UK government agencies and can be considered as self-contained labour markets and economically relevant aggregates. By combining these three sources of data, we are able to investigate how self-employment 'lines up' with some of the most salient aspects of a dense entrepreneurial environment, namely firm creation and innovation.

In terms of findings, we show that our various measures of self-employment and both net and gross firm creation, as well as the incidence of innovation, are positively and significantly correlated in *urban* areas. However, this is not the case for *rural* areas. This distinction is not easily explained by differences in the sectoral composition of businesses in urban and rural TTWAs. Although our data on self-employment become thin and our analysis less precise when we distinguish between services and manufacturing across urban and rural TTWAs, we still find the same patterns: a high incidence of self-employed in urban areas in either manufacturing or services positively correlates with higher firm creation, and either process (for services) or product (for manufacturing) innovation, but this is not true for rural areas.

In order to explain these results, we exploit additional information contained in the LFS to compute proxies for lack of employment opportunities in the TTWAs. In particular, we construct variables that measure: *(i)* the incidence of underemployment; *(ii)* the incidence of temporary employment; *(iii)* the lack of full time employment. Our analysis shows that these variables significantly predict the misalignment between self-employment, and both firm creation and innovation in rural areas. However, the same is not true for urban areas where the discrepancies among the individual-level and firm-level proxies for entrepreneurship are much smaller. These findings suggest that the urban/rural divide is related to lack of employment opportunities in rural areas, which might push some workers into self-employment as a last

resort option. All in all, our results carry some important implications for the academic debate on entrepreneurship, as well as for the design of policies that promote self-employment with the aim of stimulating business creation.

In terms of relation to the existing literature, our findings are most closely linked – and in part comparable – to Hurst and Pugsley (2010) and Sanandaji (2010). Relative to Hurst and Pugsley (2010), we cannot measure small entrepreneurs' intentions to grow and innovate. We can however link the incidence of self-employment to net firm creation – related to firm survival and expansion – and innovative activities, thus looking at these issues from a similar angle. Furthermore, similar to Hurst and Pugsley (2010), we can investigate a number of reasons for choosing self-employment as reported by individuals in the LFS. This exercise reveals that, although the fraction of individuals who chose to become self-employed because of lack of employment opportunities is slightly larger in rural than urban areas, this difference is not substantial. Moreover, the incidence of self-employed workers who were driven by the desire 'to generate income' or who identified a 'business idea/new product' niche is not significantly larger in urban than in rural areas. This evidence suggests that relying on self-reported individual assessments of personal motivation and innovative behaviour to identify entrepreneurs might mask an interesting divide along the urban/rural dimension.

In relation to the study by Sanandaji (2010), we follow a similar approach by comparing the spatial distribution of self-employed and entrepreneurs. However, our work has the advantage of focusing on one single country, thus abstracting from problems with cross-country differences in institutions and culture, as well as legal and taxation systems. Moreover, our work uses direct proxies for the most productive aspects of entrepreneurship, namely firm creation and innovation. Conversely, cross-country differences in the incidence of billionaires might partly reflect differences in taxation systems – e.g. the 50%[2] top income-tax rate discouraging business expansion in Great Britain – rather than true disparities in entrepreneurial density. Finally, our measures for firm creation are better proxies for entrepreneurship as opposed to 'entrepreneurial stardom'. Previous research shows that the density of all businesses – including *small* ones – is an important force determining agglomeration economies (Ellison et al., 2010, Glaeser, 2009 and Glaeser and Kerr, 2010) and spatial differences in economic performance. Nevertheless, if the aim of public policy was to promote a handful of very successful business ventures, the evidence in Sanandaji (2010) would be more relevant than ours.

The remainder of the paper is structured as follows. In the next section, we describe the data that we use in more details, while in Section 3 we provide general descriptive statistics. In Section 4, we present our analysis on the relation between self-employment, firm creation and

[2] As announced in the March 2012 UK Budget, the top tax rate will be reduced from 50% to 45% in April 2013.

innovation. Following that, in Section 5, we investigate some factors that could explain the urban/rural divide. In Section 6, we conclude.

2. Data construction

In this section, we describe the data that we use to address the questions highlighted here above, and discuss how we construct several measures of self-employment, firm creation and firm innovation which we then aggregate at the Travel-to-Work Area (TTWA) level.

2.1 UK Labour Force Survey (LFS)

The UK Labour Force Survey (LFS) is a quarterly representative survey of households living at private addresses in the United Kingdom conducted by the Office for National Statistics (ONS). The main purpose of the LFS is to collect information about individuals' labour market experiences to derive information that can be used to evaluate labour market policies and dynamics. The LFS data collection started in 1973 as a biannual survey. It then moved to a yearly basis and finally to the current quarterly structure in 1992. For our analysis, we use the years between 1995 and 2009, and focus on the Spring quarter since this is the part of the survey where the richest and most consistent information is available.

Although the seasonal division of the LFS (Spring, Summer, Autumn and Winter) moved to a calendar partitioning in 2006 (January-March, April-June, July-September, October-December), the ONS provides LFS data converted in calendar quarters covering the whole period 1995-2009. In many instances, however, the ONS conversion was not complete and a number of variables went partly missing. Using the seasonal quarters of the LFS prior to 2006, we were predominantly able to reconstruct information in a consistent manner and retain the calendar partitioning throughout. During the remainder of the paper, we will refer to the quarter April-June as the Spring quarter.

Each Spring quarter of the LFS contains between 64,000 (earlier years) and 52,000 (later years) households, equivalent to about 155,000 (earlier years) and 120,000 (later years) individuals. In our analysis, we only focus on people aged between 16 and 65. Additionally, since we are interested in individuals' decision to either work as employee or start their own business by becoming self-employed, we only focus on people in one of these two categories, and exclude individuals who are unemployed or not in the labour force. Since we are selecting individuals at the end of compulsory schooling (16) and before retirement age (65), we are not overly concerned by the latter restriction. However, we will come back to this point when discussing our findings. Next, in order to assign each individual to a TTWA, we retain individuals living in England, Scotland and Wales (LFS data for Northern Ireland have poor

coverage), and with a valid geographical identifier (ward of residence, roughly equivalent to a US census tract). Additionally, since in parts of our analysis we use individuals' background characteristics, we also select individuals with non-missing information on: *(i)* gender, age and ethnicity; *(ii)* marital status, household size and number of children; *(iii)* educational qualifications; *(iv)* housing tenure status. We also retain information on whether individuals work full-time or part-time, and on whether individuals hold a second job. A detailed list of control variables with descriptive statistics for information aggregated at the TTWA level is provided in Appendix Table 2. Finally, following previous work in the literature (Glaeser, 2009; and Glaeser and Kerr, 2010), we exclude individuals working in one of the following sectors: Agriculture; Hunting and Forestry; Fishing; Public Utilities (Electricity, Gas and Water); Public Administration and Defence; Private Households with Employees; and Extra-Territorial Organization and Bodies. These sectors either contain negligible amounts of entrepreneurship (e.g., public administration) or are characterised by high self-employment rates dictated by sector specific features, and not necessarily indicative of dense entrepreneurial environments (e.g., agriculture and fishing).

These restrictions leave us with a set of approximately 700,000 individuals, of which 540,000 and 160,000 live in urban and rural areas, respectively. Using this sample, we construct the following progressively more restrictive definitions of self-employment and entrepreneurship. To being with, we create a binary variable taking value one if the worker is self-employed and zero otherwise. We label the proxy 'All Self-Employed'. Next, using the socio-economic classification of jobs provided by the LFS (SOC-2000 at the 1-digit level)[3], we identify self-employed who are "Managers and senior officials", or work in "Professional occupations", or identify themselves as "Associate professional and technical occupations". We label this group 'Self-Employed in Managerial Professions' or 'Managers'. Our next definition exploits answers to the question: "Do you own the business or have a controlling interest in the company you work for?". Using this detail, we construct a proxy labeled 'Self-employed Owning/Controlling Business' or 'Owners'. Note that this question was asked to both self-employed and dependent workers in managerial and decisional positions every year up to 2004, but then every two years from 2006 onwards. Because of this limitation, the variable 'Owners' is constructed using about 400,000 individuals. Finally, we create another definition of entrepreneurs which exploits information about methods of payments for self-employed individuals. In particular, we create a proxy including individuals who report that they: "Are a sole director of their own limited business"; "Run a business or a professional practice"; "Are a

[3] LFS provides two socio-economic classifications of jobs: SOC-1990 before 2001 and SOC-2000 from 2001 onwards. At the 1-digit level, SOC-1990 and SOC-2000 are equivalent.

partner in a business or a professional practice"; "Work for themselves". On the other hand, the definition excludes self-employed individuals who are "Paid a salary or wage by an agency"; are "Sub-contractors"; or are "Doing free-lance work" plus another residual category ("None of the above"). We label this set 'Self-employed, No Freelance/Subcontractors' or 'Real Entrepreneurs'. Since information on payment methods is available from 1999 only, the variable 'Real Entrepreneurs' is derived from approximately 515,000 observations. Descriptive statistics for these variables aggregated at the TTWA level will be presented in Section 3.

2.2 The Business Structure Database (BSD)

The second dataset we exploit for our analysis is the Business Structure Database (BSD) over the period 1997 to 2008. The data is derived from the Inter-Departmental Business Register (IDBR), which consists of administrative data collected for revenues and taxation purposes. Any business liable for value-added taxation and/or with at least one employee registered for tax collection will appear on the IDBR. To complement this information, the IDBR is supplemented with data from the ONS business surveys. For 2004, the ONS estimated that the businesses listed on the IDBR accounted for almost 99 per cent of economic activity in the UK, with only very small businesses left out.

The reporting period is generally the end of the financial year (April), and the data is structured into enterprises and local units. An enterprise is the overall business organisation, whereas the local unit can be thought of as a plant or firm. Note that in the remainder of the paper, we will use these two words interchangeably. In approximately 70 per cent of the cases, enterprises only have one local unit, with the remaining 30 per cent of the cases representing enterprises with multiple local units. These include large organisations, such as banks and supermarkets, but also small multi-office consultancies and other services, as well as smaller manufacturing enterprises. For each local unit, data is available on employment, industrial activity based on the 2003 Standard Industrial Classification (SIC), year of birth (start-up date) and death (termination date), as well as postcodes. Using the latter detail, we assign each local unit active in England, Wales and Scotland to a Travel-to-Work Area (TTWA). Once again, we neglect Northern Ireland because of poor data coverage.

The initial raw data includes approximately three million local units every year. However, in order to use the data for our analysis, we carry out a series of checks and drop a number of units. First, we investigate the consistency of opening and closing dates of BSD units with their actual existence in the dataset, and drop a limited number of anomalous cases. Looking at this information, we also reclassify some units that seem to disappear and re-appear in the data with no apparent death and birth, and minimise losses of observations. However, we drop cases

where we identify establishments opening/closing in a specific year, disappearing/reappearing in a subsequent year only to open/close again in a subsequent wave. Stated differently, we only count firms' birth and death once. This approach follows Glaeser and Kerr (2010).

Next, we check the consistency of units' postcodes and sectors of activity across adjacent years, drop cases with missing information and reclassify some discrepancies over subsequent years. Furthermore, we drop active units with zero employment (this figure includes the owners/managers of the establishment, so it cannot be zero for an active unit) and postcodes that include an anomalous number of units in the same industrial sector (i.e. postcodes above the 95th percentile of the distribution of units per postcode and 3-digit SIC industry). Finally, we follow the approach used for the LFS, and exclude firms operating in one of the following sectors: Agriculture; Hunting and Forestry; Fishing; Public Utilities (Electricity, Gas and Water); Public Administration and Defence; Private Households with Employees; and Extra-Territorial Organization and Bodies.

After applying these restrictions, our dataset comprises of around 2.4 million plants annually over 12 years, which we use to derive the following proxies for gross and net firm creation. To begin with, we identify in each pair of adjacent years: *(i)* firms that were not present in wave$_{t-1}$, but are present in wave$_t$. This group identifies the number of firms that were created between period *t-1* and *t*; and *(ii)* firms that were present in wave$_{t-1}$, but are not present in wave$_t$. This group identifies the number of firms that were destroyed between period *t-1* and *t*. We can then count the number of establishments in *(i)* and *(ii)* at the TTWA level, and construct some proxies for the local intensity of business creation. In particular, using information from the first group, we can estimate the amount of *gross* firm creation in a given TTWA. Conversely, using data from both *(i)* and *(ii)*, and subtracting the number of establishments destroyed from the number of establishments created, we can proxy for the rate of *net* firm creation in a given TTWA. Note that when we aggregate our data at the TTWA level, we express net and gross firm creation as a share of the average number of firms existing at time *t-1* and *t*. More details are provided in Section 3, where we discuss the descriptive statistics of our proxies.

2.3 Community Innovation Survey (CIS)

The last dataset we use for our analysis is the Community Innovation Survey (CIS) for the years 2001, 2005, 2007 and 2009. The CIS is part of a broader European Community data collection effort and gathers firm level-data on innovation and related activities. The UK survey is structured in a way that makes the sample representative of small, medium and large businesses, across all regions and sectors. This cross-sector, cross-size sample is important in reducing biases included in other samples that track innovation, e.g. patent-level data that focus on highly

innovative sectors – but on manufacturing only and/or on firms of a larger size. It is worth noting that the survey is backed by the Department for Business, Skills and Innovation (BIS) and seen as an important tool for "measuring the level of innovation activity in the UK and identifying where policy might be best targeted" (ONS, 2011).

The original samples of the CIS varied between approximately 8000 firms in 2001 to approximately 15,000 firms in the following waves. This sample includes units belonging to both single and multi-plant enterprises, and contains a smaller panel component which we do not exploit in our work. One drawback of the CIS is that it does not directly include detailed geographical identifiers of the firms sampled. This is however a crucial piece of information in order to conduct our analysis, and compare rates of self-employment, business creation and innovation at the TTWA level. In order to recover this information, we use a unique firm identifier that allows us to match CIS firms to the BSD data described here above. BSD data contains full information on the postcode where firms' activities take place which allows us to assign firms in CIS to a TTWA. Given the quality of the BSD administrative data, the attrition from this matching is very small (below 1% in all years available).

However, the identifiers and the information contained in the CIS refer to the enterprise, as opposed to the local plant. This means that when we match data from the CIS to the BSD, we assign the same information about innovative behaviour to all plants belonging to a single enterprise since it is impossible to identify the exact local plant where the innovation actually took place.[4] While this is not ideal, we believe this procedure is better than geographically assigning information based on the location of the headquarters of the enterprise. As a robustness check, we tried to only match single-plant firms in the CIS with corresponding BSD firms, keeping approximately 75% of the sample. Reassuringly, we find that the correlation between our measures of innovation at the TTWAs obtained either using the sample of single-plant firms only or both single- and multi-plant firms is sufficiently high – at around 0.40/0.50 – and that our main findings are not affected when considering the innovative behaviour of single-plant firms only. We will return to this point in Section 4.1.

Note that following the approach used for both the LFS and BSD, we exclude firms registered as operating in one of the following sectors: Agriculture; Hunting and Forestry; Fishing; Public Utilities (Electricity, Gas and Water); Public Administration and Defence; Private Households with Employees; and Extra-Territorial Organization and Bodies.

[4] One further complication in matching the CIS to the BSD is due to the fact that approximately 2-3% of the CIS enterprises have more than one reporting unit. A reporting unit is a statistical entity created by the ONS for data collecting purposes, and stands in between the local plant and the enterprise. In the analysis that follows, we drop these enterprises since it is unclear how to assign several reporting units to the various plants of the same enterprise.

In order to construct our proxies for innovative behaviour, we use firms' answers to the following questions: *(i)* "During the three year period (prior to the survey), did your enterprise introduce any technologically new or significantly improved products (goods or services) which were new to your firm?"; and *(ii)* "During the three year period (prior to the survey), did your enterprise introduce any new or significantly improved processes for producing or supplying products (goods or services) which were new to your industry?". To begin with, we combine answers from both questions to create a dependent variable taking value one if the firm created either an innovative product (positive answer to question *(i)*) or an innovative process (positive answer to question *(ii)*). We label this variable 'Innovation'. We then create two further proxies where we separately consider whether the company successfully engaged in either product or process innovation. These variables are labelled 'Product Innovation' and 'Process Innovation'. In order to use this information in our analysis, we collapse the data at the TTWA level so that our indicators can be interpreted as fractions of innovative firms in a given area. Descriptive statistics for these variables are discussed in Section 3.

2.3 Travel-to-work-areas (TTWAs) and other geographical details

The level of geographical aggregation that we use in our analysis is the Travel-to-Work Area (TTWA). TTWAs are a statistical tool devised by UK government agencies to encompass areas within which the population commutes for employment and business purposes. Stated differently, TTWAs are functional (as opposed to administrative) geographical units that can be considered as self-contained labour markets and economically relevant aggregates.

In more detail, TTWAs are groups of wards for which at least 75% of the resident economically active population works in the area, and for which at least 75% of individuals working in the area live in the area. As from 2007, there were 243 TTWAs within the United Kingdom. In our analysis, we only focus on England, Scotland and Wales. Moreover, we follow Gibbons *et al*. (2010) and re-aggregate some areas so that our final partition splits Great Britain into 158 local economic areas, of which 79 are single urban TTWAs, and 79 are rural areas created by combining TTWAs. This classification was obtained after aggregating some contiguous rural areas with low counts of employed workers as identified by the NES/ASHE data. More details can be found in Gibbons *et al.* (2010). Differently from Gibbons *et al.* (2010), we distinguish between Inner London and Outer London by splitting the London TTWA in two parts. We follow this approach because we believe entrepreneurship, density of start-ups and

innovation might differ between the core of London and its periphery. This presumption is partly borne out by the data.[5]

Expectedly, urban TTWAs have substantially higher population density than rural areas. The average/median number of people per squared kilometre is 800.2/528.6 and 180.6/144.3 in urban and rural areas respectively, with standard deviations of 1073.7 and 140.8. Note that although there is more variation within the urban group, the 25th percentile of the population density distribution in urban areas (at 315.0) still lies above the 75th percentile of the density distribution in rural areas (at 250.7). These comparisons are not significantly affected if we disregard London: the average and median urban population density are still 651.8 and 510.7 respectively, with a standard deviation of 424.9.

Before moving on, it should be noted that data from the BSD and the CIS are assigned to TTWAs based on the postcode at which firms are active. On the other hand, individuals in the LFS are assigned to TTWAs using postcodes of residence since detailed information about their place of work is not available. We do not see this as a major drawback. First, given the way in which TTWAs are constructed (i.e. 75% of the resident population work in the area), the TTWA of an individual's residence is also likely to be his/her TTWA of employment. Moreover, this problem is further attenuated by the fact that we concentrate on the geographical distribution of self-employed individuals. Previous research has shown that entrepreneurs tend to be local and set up their business in the location where they were born and grew up (Michelacci and Silva, 2007). Finally, we shed directly some light on this issue by using LFS data at the individual level to run some regressions that estimate the probability that a self-employed individual: *(i)* works in the Local Authority/District (LAD) where he/she lives; *(ii)* works from the home address or uses his/her home as the headquarters of his/her activities.[6] Note that the sample only includes working individuals, i.e. employees and self-employed. Our evidence is presented in Appendix Table 1. Columns (1) to (4) focus on individuals in urban TTWAs, whereas Columns (5) to (8) concentrate on rural areas. Note also that even columns contain a detailed set of controls, including individual characteristics, sectors of employment and TTWA fixed effects. More details are provided in the note to the table.

Across all columns and irrespective of the definition of entrepreneur, we find that self-employed individuals are significantly more likely than employees to work in the LAD where they live or to use their home as their workplace or the headquarters of their business. These differences are economically sizeable. Entrepreneurs are between 20 and 30 percentage points –

[5] Another difference between our classification and the one in Gibbons *et al.* (2010) is that we classify Colchester and Clacton as rural. This does not affect our results and leaves us with equal numbers in the two sub-sets.

[6] There are approximately 400 Local Authorities and Districts in Great Britain although this number changes slightly over time.

or 40% to 50% – more likely to work in the LAD of residence than employees, both in urban and rural areas. Similarly, self-employed in both urban and rural areas are between five and ten times more likely than employees to work from home or use it as their headquarters. These results are not surprising, and suggest that assigning self-employed workers to areas on the basis of their TTWA of residence (as opposed to TTWA of work) does not introduce an important bias in the measurement of the density of entrepreneurial activities based on self-employed individuals in the LFS.

3. Self-employment, business creation and innovation: some descriptive statistics

In this section, we present descriptive statistics of the various proxies for entrepreneurship created using the three datasets described here above. Our findings are presented separately for urban and rural TTWAs, and they are depicted in Figures 1.a-1.b and tabulated in Table 1.

3.1 Self-employment in urban and rural TTWAs: main facts

We begin our discussion by presenting descriptive statistics for the proxies for entrepreneurship obtained from the LFS using various definitions of self-employed workers. The figures in the top panel of Table 1 (All Self-Employed) confirm some stylised facts about self-employment in Great Britain previously obtained using the LFS (Blanchflower and Shadforth, 2007) or other data (e.g. Bracke *et al*., 2012 using BHPS): over the period spanning the mid-1990s to the late 2000s, approximately 12%-13% of workers are self-employed, with this number not being substantially different in urban and rural areas. However, there is quite a significant amount of heterogeneity across TTWAs, as shown in Columns (3) and (6). The coefficient of variation for the share of self-employed across TTWAs is approximately 19% and 24% in urban and rural areas, respectively. While the figures for rural areas might portray more variation than there actually is because of small numbers, this urban/rural ranking is repeated when we look at entrepreneurship measures derived from firm-level data, suggesting that there might be more clustering of high and low entrepreneurship hot-spots in rural areas.

Similarly, there is quite a substantial variation in the share of self-employed individuals when we consider manufacturing and services separately. First, we find that more individuals are self-employed in services than in manufacturing. This is hardly surprising and is consistent with previous evidence (Blanchflower and Shadforth, 2007; Hurst and Pugsley, 2010 and Glaeser, 2009). However, there is substantially more variation across TTWAs in the share of self-employed in manufacturing than in services. The coefficient of variation in the share of self-employed is between 47% and 52% for manufacturing, and between 16% and 22% in services. One important caveat in interpreting these patterns is that the amount of variation in the

incidence of self-employment in manufacturing might be inflated by the smaller number of individuals used to aggregate information at the TTWA level. Nevertheless, this finding is quite stark and partly intuitive: the geographical distribution of activities in services – including retail trade, entertainment, professional and personal services – arguably follows more the distribution of the population than the distribution of manufacturing plants, which is instead influenced to a larger extent by agglomeration forces (Dopeso-Fernandez, 2010).

In the next three panels of Table 1, we present descriptive statistics of our alternative proxies for entrepreneurial workers, namely 'Self-Employed in Managerial Professions'; 'Self-Employed Owning/Controlling the Business' and 'Real Entrepreneurs'. Expectedly, these alternative definitions identify fewer entrepreneurs, but there are no marked differences between urban and rural areas. Once again, the amount of variation across TTWAs is substantial and more pronounced in rural areas – between 24% and 29% – than in urban areas – at 19% and 21% for 'Real Entrepreneurs' and 'Owners'. This variation further climbs to 28% for 'Managers' in urban areas. This finding is partly driven by Inner London where the incidence of self-employed in managerial professions is substantially larger (approximately five percentage points) than the urban sample mean.

Note also that there are significantly less entrepreneurs across all these three categories when focusing on manufacturing as opposed to services, and that this reduction is very pronounced for 'Managers', where the share of self-employed in manufacturing falls to 2.2% and 2.6% in urban and rural areas, respectively. The corresponding shares of self-employed in services are three-time as large at 6.7% and 7.7%. Similar, but less pronounced patterns can be found for the other two definitions. For example, the share of 'Real Entrepreneurs' in manufacturing and services in urban and rural areas are respectively: 4.8% and 9.8%; and 6.2% and 11.1%. Note finally that for these definitions too we document more variation in manufacturing than in services, although this might be partly due to the thinness of our data.

A graphical impression of these patterns is presented in Figures 1.a and 1.b, where we map the density of our four proxies for entrepreneurial workers across urban and rural TTWAs. The urban-area plots (Figure 1.a) confirm the common perception that London and the South-East are more entrepreneurial. Although this is true for all definitions, the pattern is more marked for the more restrictive definitions of entrepreneurs, in particular for the definition that excludes freelancers, subcontractors and agency workers ('Real Entrepreneurs'). Some expected patterns also emerge when focusing on the rural maps (Figure 1.b). These show a high density of self-employment along the South-West coast, in Cornwall and in parts of Wales and Scotland, which is partly explained by the tourist industry. However, there is also a high incidence of self-

employment in some northern areas along the corridor running east to west, and north of Hull, York, Manchester and Liverpool.

Another finding that emerges from these plots is that the more or less restrictive definitions of entrepreneurial individuals tend to highlight hot-spots of entrepreneurship in similar areas. In fact, the correlations between the various measures of self-employment are high and vary between approximately 0.83 (between 'Owners' and 'Managers') and 0.97 (between 'All Self-Employed' and 'Real Entrepreneurs') in both urban and rural areas.

3.2 Firm creation and innovation across TTWAs: main facts

In this section, we discuss the main properties of the proxies for entrepreneurship that we have constructed by aggregating data on firm birth and death from the BSD, and business innovative activities from the CIS. Descriptive statistics are reported in the bottom three panels of Table 1.

Starting with firm entry or *gross* firm creation, approximately 14% new firms are created every year in both urban and rural areas. Nevertheless, the yearly rate of *net* firm creation is around 0.5% in both urban and rural areas, implying that at every point in time nearly as many firms enter as exit the market.[7] Once again, the variation across TTWAs is very significant: the coefficients of variation for the net share of firm creation for urban and rural areas are respectively 89% and 93%. The larger dispersion in rural areas is in line with the findings discussed above for the LFS. We find a similar pattern when looking at the gross share of firm creation, although the extent of variation is much smaller at 4.3% and 6.3% in urban and rural areas respectively. This suggests that a substantial part of the geographical differences in terms of firm density is explained by survival rates, and firms that enter the market only to exit one year later (i.e. churning).

Next, we partition our measures of firm creation between services and manufacturing. For services, we find that in both urban and rural areas, on average between 1995 and 2008, the number of establishments has been expanding. The gross rate of firm creation was in the order of 14% every year, while the net rate was 0.6%-0.7%. On the other hand, manufacturing activities have been shrinking, and more markedly so in urban areas. Although the gross rate of manufacturing firm creation in both urban and rural areas was approximately 10% per year, the net rate was -2.4% in urban areas and -1.5% in rural areas. Furthermore, there is more variation across TTWAs in net firm creation in the service industries than in manufacturing, while the opposite is true for gross firm creation. The larger extent of spatial variation in services than

[7] Note that, considering all sectors, the correlation between *net* and *gross* firm creation is 0.65 and 0.67 in urban and rural TTWAs respectively.

manufacturing for net firm creation is in contrast with the results we obtained using the LFS, and might be partly explained by the general downward trend in manufacturing. Conversely, the more substantial variation across TTWAs in manufacturing than in services that we obtain looking at gross rates of firm creation is in line with the results documented using the LFS.

Finally, in the bottom panel of the Table 1, we report descriptive statistics for the fraction of firms that innovate in the TTWA. This proxy bundles together both process innovation and product innovation. Looking through the columns, the figures show that firms in urban areas innovate slightly more than those in rural areas – with shares of innovative firms at 30.4% and 29.5% respectively – but these differences are not marked. Furthermore, there is more dispersion in innovative activities across rural areas (coefficient of variation 12%) than urban areas (8.16%). This is in line with the picture presented so far. Next, we also find that more firms innovate in manufacturing (approximately 46% in both rural and urban areas) than in services (around 26.8% and 28.4% in rural and urban areas), and that this is true even if we focus on *process* innovation for service industries and *product* innovation for manufacturing. The share of service firms creating new processes is approximately 20% in both urban and rural areas, while the incidence of manufacturing firms inventing new products is around 45% across the board. Finally, there is more variation across TTWAs in manufacturing than in services. Although these figures might be affected by fewer manufacturing firms than service firms in the CIS (due to the representative nature of the survey), this pattern follows the trend documented for our other proxies.

3.3 The sectoral distribution of self-employed workers and firms

To conclude this section, we present descriptive statistics for the sectoral distribution of self-employed individuals in the LFS and the distribution of firms in the BSD data. Note that in this case, we append all plants active in the BSD in the various years to add up to about 29 million observations (or 2.4 million observations a year). Our findings are presented in Table 2. The top panel reports figures for urban TTWAs, while the bottom panel presents statistics for rural areas. Note also that we re-grouped sectors to partly match Glaeser (2009) and Hurst and Pugsley (2010). More details are provided in the note to the table.

Starting with the urban sectors, the overall impression is that, irrespective of the self-employment definition we consider, the match between the sectoral distribution of LFS entrepreneurs and firms in the BSD is reasonably good. The biggest discrepancies are concentrated in Construction; Entertainment; and Health Care Services. The percentage of BSD units in Construction is 9.82%, while this figure is around 20% for self-employed workers in the LFS, except for ‘Managers’ where it drops to 4%. Conversely, the incidence of ‘Managers’ in

Entertainment and Health Care Services – at 11% and 10% – over-represent the corresponding shares from the BSD – at 4% and 5% respectively. However, these discrepancies are attenuated when using the other three definitions of self-employed entrepreneurs. Finally, the incidence of Wholesale Trade; Finance, Insurance and Real Estate (FIRE); and Accommodation/Food Services is higher in the BSD than in the LFS. However, the discrepancies within the latter two sectors (FIRE and Accommodation/Food) are smaller when considering 'Managers'.

As for rural areas, differences between the BSD and LFS are similar to those detected in the urban sample. However, there is a higher incidence of self-employed (LFS) and firms (BSD) in Construction; Retail Trade; and Accommodation and Food Services; and a smaller incidence of Professional Services. Some of these urban/rural differences are more pronounced when considering some of the self-employment definitions (e.g. 'Managers'), but broadly speaking similar patterns emerge from the BSD and LFS.

Before moving on, it is instructive to compare the sectoral incidence of the UK self-employed with figures provided in Glaeser (2009) and Hurst and Pugsley (2010) for the US. Glaeser (2009) tabulates the incidence of self-employed workers in non-agricultural sectors. Our figures are broadly comparable to his, although we tend to over-sample self-employed in Construction (except for 'Managers') and have slightly more self-employment in High-Tech Manufacturing, Accommodation and Food, and Health services. Conversely, we have less self-employed in Low-Tech Manufacturing. Note that we also have a larger group of Professional and FIRE self-employed, broadly speaking corresponding to Glaeser's high-skill Information Services. Relative to Hurst and Pugsley (2010), LFS self-employed workers tend to feature more prominently in Construction (except for 'Managers'), but also in allegedly more entrepreneurial sectors such as manufacturing – in particular High Tech Manufacturing – and Professional services. There are also some discrepancies in the share of LFS self-employed workers in Transportation and Warehouse (more in the LFS, although this group further includes Communication Services in our data), and FIRE (less in the LFS). However, these discrepancies are much attenuated when focussing on 'Managers'.

By and large, these comparisons reveal that the sectoral distribution of self-employed and small businesses in the US and the UK is remarkably similar. We believe this suggests that our main results should extrapolate to US self-employed and start-ups.

4. How good a proxy for entrepreneurship? The link between self-employment, firm creation and innovation

4.1 Urban and rural areas: evidence pooling all sectors

In this section, we exploit information from the three datasets discussed above combined at the TTWA level to investigate whether proxies for entrepreneurship based on individual self-employment are well aligned with business start-up rates and innovation density.

Our results are presented graphically by cross-plotting shares of self-employed individuals in the different TTWAs according to our various definitions against rates of firm creation and innovation. Further, to assess the significance of the relations depicted in the graphs, we also report results from simple regressions at the TTWA of either firm creation rates or the incidence of innovation on the share of self-employment. Results are reported in the various panels of our figures that are organized as follows: Figures 2.a-2.d focus on urban areas, whereas Figures 3.a-3.d focus on rural areas. Figures 2.a and 3.a focus on 'All Self-Employed'; Figures 2.b and 3.b focus on 'Managers'; Figures 2.c and 3.c focus on 'Owners'; and finally Figures 2.d and 3.d focus on 'Real Entrepreneurs'.

Starting with the urban plots, the top left panels of Figures 2.a to 2.d show the alignment of average TTWA self-employment rates up to 2002 and after 2002. These plots show that roughly speaking TTWAs with more entrepreneurs up to 2002 remained more entrepreneurial in the subsequent years. Interestingly, these graphs also confirm that the South East and in particular London and the surrounding areas are highly entrepreneurial: Brighton, Tunbridge Wells, Guilford, Inner and Outer London, Cambridge, Oxford and Slough always feature at the top. Unsurprisingly, Inner London ranks very high in terms of density of managerial entrepreneurs. Conversely, Dudley, Dundee, Bradford, Glasgow and Newport always rank near the bottom.

Next, moving clockwise, we present the alignment of self-employment with: *(i)* net firm creation (top right panel); *(ii)* gross firm creation (bottom right panel); *(iii)* share of innovative firms (bottom left panel). All variables have been demeaned so that the scale on the axis is in deviations from sample means. The numbers at the bottom of each panel report regression coefficients of the demeaned variables on the vertical axis on the horizontal axis, and associated heteroskedasticity-robust standard errors.

The overall pattern presents a positive story: rates of self-employment are reasonably well aligned with firm creation and innovation intensity. The relation between the incidence of self-employed workers and net firm creation is always positive, and the regression coefficients reveal a significantly positive association for all self-employment proxies with t-statistics in the range of 3.5-4. For example, the regression coefficient of net firm creation on 'Managers' is 0.109 with a standard error (s.e.) of 0.030, whereas the coefficient for 'Owners' is 0.104 with an s.e. of 0.026. When looking at gross rates of firm creation, we find that the association with self-employment remains positive, although overall levels of significance are attenuated. For example, the regression coefficient of gross firm entry rates on 'Managers' is 0.105 with an s.e.

of 0.060, significant at the 10% level. Similarly, the relation between gross firm creation and 'Real Entrepreneurs' is significant at the 10% level, with a coefficient of 0.060 (s.e. 0.035). These findings are partly explained by a more compressed distribution in the rates of *gross* firm creation across TTWAs compared to the variation in *net* firm creation (as noted in Section 3.2). This suggests that there must be a large amount of heterogeneity across TTWAs in the share of firms that survive every year, and that the incidence of firm destruction in urban areas must be negatively correlated with the density of self-employment.

Finally, in the bottom left plots of Figures 2.a-2.d, we present the relation between self-employment and innovation. Across all panels, we find a positive relation between proxies for entrepreneurship derived from self-employment data and firm-level information on the intensity of innovation. This association is less strong for 'All Self-Employed' – with a coefficient of 0.266 and an s.e. of 0.134 (t-statistics of 1.98) – than for the other self-employment measures. For example, using 'Managers' we find a coefficient of 0.464 and an s.e. of 0.193 (t-statistics of 2.40), while focussing on 'Real Entrepreneurs' gives a coefficient of 0.373 associated to an s.e. of 0.155 (t-statistic of 2.41). These results suggest that the more restrictive proxies of entrepreneurship obtained by narrowing our self-employment definition help identify more productive and innovative entrepreneurs.

The positive findings for urban TTWAs are completely reversed when focussing on rural areas. Our results are reported in Figures 3.a-3.d, which follow the structure of Figures 2.a-2.d. Starting from the top left plots, we see that the intensity of self-employment before 2002 in rural areas is still reasonably well in-line with the share of self-employment after that date, although the alignment is less precise than for urban areas. Cornwall, Devon, Kendal, parts of Wales and Yorkshire feature at the top of the rankings, whereas Scottish TTWAs tend to be at the bottom of the self-employment distribution. This is true irrespective of the definition we consider.

More importantly, the share of self-employment in rural TTWAs is *not* positively and significantly associated with the share of net and gross firm creation. Starting with the former, although there is a positive association between net firm creation and our proxies for self-employment, this relation is much flatter than for the urban sample and not significant at conventional levels. For example, the coefficient of the regression of the net share of firm creation on 'Managers' is 0.019 (s.e. 0.029), while this figure for 'Real Entrepreneurs' is 0.015 (s.e. 0.020). As for the gross share of firm creation, this is *negatively* linked to the share of self-employed workers in the TTWA. While this negative link is not statistically significant, it is more pronounced for more restrictive categories of self-employed: for 'Managers' and 'Real Entrepreneurs' we find regression coefficients of -0.082 (s.e. 0.051) and -0.036 (s.e. 0.035), respectively.

Similarly, the bottom left panels of Figures 3.a-3.d reveal that the share of self-employed workers in a TTWA is not positively associated with innovative activities. While for three out of four measures of self-employment the relation with innovation remains positive, this link is never significant and smaller than previously documented. For 'Managers', the regression of the share of innovative firms in a TTWA on the fraction of self-employed workers yields a coefficient of 0.130 with an s.e. 0.193; while for 'Real Entrepreneur', these numbers are 0.071 and 0.133. Finally, the relation between 'Owners' and innovation turns negative (not significant) with a regression coefficient of -0.068 (s.e. 0.172).

Note that at the bottom of each panel we also report the R-squared from our simple TTWA-level regressions in urban and rural areas. Across all specifications, R-squared are substantially higher for the urban sample than for the rural one. For instance, the incidence of self-employment explains between 11% and 16% of the overall variation in net firm creation in urban areas. The corresponding measure for rural areas ranges between 0.5% and 2%. Similarly, urban self-employment rates explain between 6% and 9% of the spatial variation in innovation activities. The corresponding figures for the rural sample are lower, at 0.1%-0.5%.

Finally, we checked whether these patterns are affected by the inclusion/exclusion of multi-plant enterprises from the BSD and the CIS, and found that this is not the case. When excluding multi-plant firms from the BSD, the relation between net firm creation and either 'Managers' or 'Real Entrepreneurs' remains positive in urban areas, with regression coefficients of 0.053 (s.e. 0.034) and 0.032 (0.026), respectively. Conversely, these figures turn negative in rural areas, respectively at -0.028 (s.e. 0.035) and -0.016 (0.023). Similarly, if we consider the link between self-employment and innovation excluding multi-plant firms from the CIS, we find positive results for urban areas but negative ones for rural TTWA. For example, regressing the share of innovative firms on either self-employed 'Managers' or 'Real Entrepreneurs' yields coefficients of 0.405 (s.e. 0.370) and 0.542 (s.e. 0.269), respectively. These numbers for rural TTWAs turn markedly negative at -0.664 (s.e. 0.455) and -0.288 (s.e. 0.310).

All in all, these findings clearly suggest that measures of self-employment derived from individual-level data act as good proxies for entrepreneurship across urban areas. However, the same measures applied to rural areas would give a distorted picture of the spatial distribution of entrepreneurial activities.

4.2 Sectoral heterogeneity: manufacturing versus services

As already discussed, the share of firms in manufacturing sectors has been declining steadily in Great Britain during the period 1997-2008, but this decline was more marked in urban areas than in rural areas, leaving more room for services to flourish. It might also be argued that self-

employment is a better proxy for entrepreneurship in service sectors, where individuals leading consultancies, professional and IT services, working in finance, or setting up wholesale and retail chains might still qualify as self-employed, and yet be highly entrepreneurial. On the other hand, entrepreneurs who founded a larger manufacturing plant – despite being innovative and entrepreneurial – might not identify themselves as self-employed. If this was the case, the lack of alignment between the share of self-employed entrepreneurs and both firm creation and innovative activities in rural areas might be related to sectoral considerations.

In order to explore this issue, in Figures 4 and 5 we repeat the analysis carried out in Section 4.1, but considering service and manufacturing industries separately. Note that in this analysis we only consider 'Real Entrepreneurs' as our proxy for entrepreneurship derived from self-employment data. Results for the other definitions were similar and are available upon request. Further, note that when calculating the share of self-employed workers in urban and rural areas, and separately for services and manufacturing, our individual-level data becomes thin. This is particularly true for manufacturing in rural areas, because fewer people live in rural areas and work in manufacturing (the number of individuals working in rural manufacturing is about 30,000, climbing to 90,000 for urban manufacturing). Therefore, we regard the evidence in this section as more 'noisy' than the findings discussed above.

Figure 4 concentrates on individuals and firms operating in service industries. The left panels of the figure refer to individuals and firms located in urban areas, whereas the right panels concentrate on rural areas. The three different panels present evidence on the link between self-employment and: *(i)* net firm creation (top panel); *(ii)* gross firm entry (middle panel); *(iii)* innovation (bottom panel). Notice that since we are focussing on services, we concentrate on *process* innovation.

Starting with the urban graphs, we note that the share of 'Real Entrepreneurs' is positively aligned with all three firm-data proxies for entrepreneurship in urban areas, although this relation is significant for net firm creation (coeff. 0.073; s.e. 0.036) and process innovation (coeff.= 0.414; s.e. of 0.118), but not for gross firm creation (coeff. 0.055; s.e. 0.045). On the other hand, the relation between 'Real Entrepreneurs' in rural areas and firm creation is flat when considering net rates of business start-ups (coeff. -0.008; s.e. 0.028), and significantly negative when focusing on gross firm creation (coeff. -0.099; s.e. 0.042). Finally, the link between self-employment rates in rural areas and the share of innovative firms is positive, but less precisely estimated and sizeable than for urban areas (coeff. 0.224; s.e. 0.081).

Next, in Figures 5, we replicate this analysis for manufacturing. Once again, the reader should bear in mind that our self-employment rates are calculated over a small number of

individuals, which makes our findings more tentative. Note also that the proxy for innovation now considers only firms that engage in *product* innovation.

Starting again with the urban panels, we find that a larger share of 'Real Entrepreneurs' in urban areas is positively associated with net firm creation (coeff. 0.074; s.e. 0.032), gross firm creation (coeff. 0.080; s.e. 0.042) and innovation (coeff. 0.695; s.e. 0.386). However, this is not the case for rural areas. The link between self-employment rates and net firm creation remains positive, but insignificant (coeff. 0.055; s.e. 0.022), while the relation between 'Real Entrepreneurs' and firm entry is negative, although insignificant at -0.003 (s.e. 0.035). Finally, the incidence of self-employment is negatively correlated with the share of innovative firm, although this relation is not significant (coeff. -0.384; s.e. 0.229).

All in all, the evidence gathered in this section suggests that the stark urban/rural divide cannot be explained by differences in the incidence of manufacturing and services in urban and rural TTWAs. In the next section, we use additional information contained in the LFS to try to explain this geographical divide.

5. What explains the urban/rural divide? Some insights

5.1 Individual reasons for choosing self-employment

In order to shed some light on the urban/rural divide documented above, in this section we start by investigating individuals' self-reported reasons for choosing self-employment. This data comes from the LFS and information was collected for self-employed individuals in 1999, 2000 and 2001, so the number of observations used to create this information is small and approximately 9,000 in the urban sample and 3,000 in the rural group.

Our findings are tabulated in Table 3. A number of reasons were listed by the LFS which we have regrouped as follows: 'Non Pecuniary Reasons'; 'Generate Income'; 'Business Idea/New Product'; 'Join Family Business'; 'Lack of Employment Opportunities'; and 'Other Reasons'. More details are reported in the note to the table. Note also that individuals could report up to four reasons for choosing self-employment, although only few did.[8] So in the table we report both the percentage of people reporting a given category in their first answer, as well as in any of their answers. The two left columns refer to urban TTWAs, while the two right columns focus on rural areas.

Between 36% and 39% of self-employed workers in urban areas reported 'Non Pecuniary Reasons' as their first reason for choosing self-employment, while these figures are slightly

[8] The percentage of people who gave multiple answers is very small. On average over the three years under consideration, the figures were as follows: about 17% gave two answers; 6% three answers; and 2% four answers.

smaller (between 35% and 37%) in the rural sample. Note that 'Non Pecuniary Reasons' include 'Wanted to be independent/experience a change' and 'Wanted better working conditions', so that it is hard to gauge whether these figures should predict that more entrepreneurial types choose self-employed in urban rather than rural areas.

Next, the percentage of urban workers who chose self-employment because they wanted more money ('Generate income') is between 6% and 7.5% when only considering the first reason given, and between 12% and 14.5% when all reasons are considered. Interestingly, these figures do not substantially vary across the various definitions of self-employment. Moreover, they do not seem to differ substantially between the urban and the rural sample. This result is somewhat puzzling given that we have documented in Section 4 that self-employed workers in urban areas seem to be more entrepreneurial than those in rural areas. A similar intuition can be obtained when focusing on the category 'Business Idea/New Product', including individuals who chose self-employment because 'The opportunity arose (capital, space and equipment was available)' or because 'They saw some demand/a market' for their business idea. Across all definitions of self-employment, the percentages of individuals who report this type of motivations as either their first reason or as any reported reason is *larger* in rural areas. This suggests that by looking at these self-reported intentions might not necessarily help pinning-down highly entrepreneurial types.

Further down, we see that around 4% of self-employed workers in urban areas report joining the family business as their first reason, with this percentage rising to 4.5%-5% if we consider all answers. These figures are slightly larger in rural areas, where they range between 5.3%-5.8% and 6.1%-6.7%. This could suggest that more self-employed workers in rural areas join low-productivity, non-innovative 'mom-and-pop businesses' – such as shops and other small retail units – or become second generation entrepreneurs – which are often less innovative and talented than their predecessors (Bertrand and Schoar, 2006 and Bertrand et al., 2008).

These results could also suggest that some individuals decide to join their family businesses because of lack of job opportunities. Further evidence on this issue can be gathered looking at the incidence of individuals who explicitly claim to have become self-employed because of lack of alternative employment. The share of self-employed out-of-necessity is between 9% and 11% considering only first answers, and 11%-13% when considering all answers in urban areas. Figures are slightly larger at 10%-12% and 12%-14% in rural areas. This suggests that lack of employment opportunities might explain some of the urban/rural divide presented above,

although these differences are small, casting further doubts on the validity of looking at individuals' own perceptions to identify entrepreneurial types.[9]

Before moving on, we compare our results with those of Hurst and Pugsley (2010) who tabulate nascent entrepreneurs' reasons for starting a business. Relative to our findings, a significantly larger share of their entrepreneurs report that they opened an activity to 'Generate income' – at 19.5%-21% considering the first reason only – or because they had a 'Business idea or created a new product' – at 28%-32%. On the other hand, the share of individuals who became entrepreneurs because of lack of other employment opportunities is smaller for Hurst and Pugsley (2010) – at 2.2%-2.6% considering the first reason only – than for us. Finally, the share of nascent entrepreneurs motivated by 'Non pecuniary reasons' is similar to ours, at approximately 35%-38% considering the first reason only.

To summarise, the results in this section suggest that UK self-employed workers are less entrepreneurial than small business owners in Hurst and Pugsley (2010), and that this finding is true for both urban and rural workers. However, the fact that we find no substantial differences between urban and rural TTWAs when looking at self-reported reasons for becoming self-employed contrasts with our findings presenting a clear urban/rural divide. This casts doubts about the information content of individuals' stated reasons for starting a business.

5.2 Lack of employment opportunities and the urban/rural divide

In this section, we explore more systematically whether more self-employed individuals in rural areas chose to start their business because of lack of employment opportunities. In order to do so, we use additional LFS information that allows us to measure the pervasiveness of under-employment and lack of employment opportunities among workers over several years (and not just self-employed for three years as in the previous section).

In order to construct our first proxy, we exploit information about employed workers who would like to work longer hours (in their current job), but are not offered the possibility to do so. Using this information, we construct the share of workers who are 'underemployed' in different urban and rural TTWAs. Next, we identify individuals whose job is not permanent, but fall in one of the following categories: seasonal; fixed-period; fixed-task; agency temping; casual type; other. We then aggregate this data at the TTWA level to measure the incidence of individuals with 'temporary employment'. Finally, we identify individuals who work part-time because they could not find full-time jobs. Using this detail, we impute the pervasiveness of 'lack of full time

[9] Note also that the fraction of individuals who chose self-employment because of 'Other reasons' is slightly larger in urban areas than in rural ones. However, it is impossible to attach any meaningful interpretation to this category.

employment' opportunities in the TTWA. Descriptive statistics for these proxies are presented in Appendix Table 3, separately for urban and rural TTWAs.

We think these measures are better than those we could obtain by aggregating self-employed workers' reasons to start a business at the TTWA level for two reasons. First and most importantly, these proxies can be calculated over a large number of individuals and over an extended time period, allowing us to reduce some of the 'noise' in our measures.[10] Secondly, we believe the questions we use to create the proxies discussed here above are more objectively presented and potentially elicit more reliable information than those asked to self-employed individuals about why they started their activity.

Nevertheless, the correlation between our three proxies (the incidence of underemployment, temporary employment, and the lack of full-time employment) and the share of self-employed workers who started their activity because of lack of employment opportunities is positive, although low (at around 0.20 in urban areas and 0.10 in rural ones). Further, when we replicate the regression analysis that follows using the latter variable, we find results that are consistent with the evidence presented in Table 4, although much weaker. These inconclusive findings can be most likely explained by the thinness of the data used to compute this variable.

In order to shed some light on the urban/rural divide, we use the three proxies discussed above to predict the differences between the incidence of self-employment and either net firm creation or firms' innovative behaviour, separately for urban and rural areas. Our results are presented in Table 4. Panel A focuses on urban TTWAs, while Panel B presents evidence for rural areas. Each cell in the table reports regression coefficients (with robust standard errors) from separate regressions, where the dependent variable is one of the following measures: *(i)* the difference between the standardized share of 'Real Entrepreneurs' and the standardized share of net firm creation (Columns 1 and 2, Panel A and B); *(ii)* the difference between the standardized share of 'Real Entrepreneurs' and the standardized share of innovative firms (Columns 3 and 4, Panel A and B). Note that the standardization is carried out within urban and rural TTWAs separately, and that using the other proxies for self-employed entrepreneurs and gross firm creation provided a similar intuition. Results are not reported for space reasons, but are available upon request.

The explanatory factors of interest are the proxies derived above, namely: *(i)* the incidence of underemployment; *(ii)* the incidence of temporary employment; and *(iii)* the incidence of lack of full-time employment. Note that these explanatory factors have been standardized too.

[10] In details, the incidence of underemployment and temporary employment are constructed aggregating information from approximately 2.8 million observations, while the pervasiveness of lack of full time-employment uses information for approximately 800,000 part-time workers. Conversely, the proxies obtained by aggregating self-employed individuals' responses would rely on only about 1350 observations.

Columns (1) and (3) do not include controls, whereas Columns (2) and (4) add characteristics aggregated from the LFS at the TTWA level. These include individuals' education levels, as well as the incidence of unemployment and inactivity.[11] More details are provided in the note to the tables.

Panel A of Table 4 reveals that our proxies for lack of employment opportunities do not have a strong relation with differences in self-employment rates and either net firm creation or innovation in urban areas: only one out of twelve regression coefficients is significant. This is not surprising given the good mapping between individual-level and firm-related proxies for entrepreneurship documented in Section 4.1.

On the other hand, as shown in Panel B of the table, lack of employment opportunities is a strong predictor of differences between the incidence of self-employment and net firm creation or firms' innovative behaviour in rural areas. Nearly all of the regression coefficients on our proxies for poor labour market conditions are statistically significant, and larger than in the urban sample (coefficients are comparable since all variables are standardized).

Finally, note that although precisely quantifying these effects is beyond the scope of this study, the R-squared of the (unconditional) regressions carried out in Columns (1) and (3) of Tables 4 are always larger for the rural sample than for the urban one. For urban TTWAs, these range from virtually zero (for the regression of the difference between the shares of 'Real Entrepreneurs' and net firm creation on 'lack of full-time employment') to 1.2% (for the regression of the difference between the shares of 'Real Entrepreneurs' and innovation on the 'lack of full-time employment'). Conversely, for rural TTWAs, they are between 13% (for the regression of the difference between the shares of 'Real Entrepreneurs' and innovative firms on the incidence of 'underemployment') and 20% (for the regression of the difference between the shares of 'Real Entrepreneurs' and net firm creation on 'lack of full-time employment').

In conclusion, our evidence suggests that the urban/rural divide documented so far is related to lack of employment opportunities in rural areas, which might push some workers into self-employment as a last-resort option. This finding carries important implications for the academic debate in the field which has widely used information on individuals' self-employment status to identify entrepreneurs and study a number of factors that promote entrepreneurship – such as skills – or hamper entrepreneurial activities – such as credit

[11] Note that local unemployment rates did not enter these regressions significantly as controls. Similarly, they were not strong predictors of TTWA differences in self-employment and other measures of entrepreneurship when used as explanatory factors unconditional on other controls. For example, local unemployment entered the regression of the difference between 'Real Entrepreneurs' and net firm creation across rural TTWAs with a coefficient of 0.060 and an associated s.e. of 0.134. The corresponding number for urban areas was -0.123 (se. 0.112).

constraints. Similarly, our results carry profound implications for public policies that promote self-employment with the aim of stimulating business creation and innovation.

5.3 Replicative and innovative entrepreneurship in urban and rural areas

The geographical divide presented above could also be rationalised by the distinction between 'replicative' and 'innovative' entrepreneurs and their differential incidence in urban and rural areas. As discussed in the Introduction, innovative entrepreneurs generate new economic opportunities by creating new products and processes. Conversely, replicative entrepreneurs respond to local demand by providing required goods and services. Stated differently, innovative entrepreneurship is predominantly a supply-side phenomenon, while replicative entrepreneurship is predominantly a demand-driven economic activity.

In order to test whether a high share of replicative entrepreneurship can explain the misalignment between self-employment rates and the incidence of net firm creation and innovation in rural areas, in the bottom two rows of Panels A and B of Tables 4, we investigate whether these differences are significantly linked to the population density of the TTWA. This is measured by the number of people recorded in the GB Census 2001 per squared kilometre. According to Baumol (2011), a larger population base should be associated with a larger incidence of replicative entrepreneurship, since higher population density results in higher demand for goods and services, including those provided by routine entrepreneurs.

Starting with urban areas, the top panel of Table 4 shows that the association between population density and the difference between the incidence of self-employment and net firm creation is positive overall, but not significant. Similarly, we find little evidence of a significant relation between population density in urban areas and the difference between the share of self-employment and the incidence of innovative firms.

Conversely, Panel B of Table 4 shows that the relation between population density in rural areas and the difference between net firm creation and the share of 'Real Entrepreneurs' is negative and significant at conventional levels. This is true irrespective of whether further controls are included in our analysis. Similarly, we find that the relation between population density and the difference in the incidence of self-employment and innovative firms is negative and significant in Column (4) when controls are added.

This evidence is hard to reconcile with the notion that some of the self-employment spells in the densest rural areas can be considered routine entrepreneurship. For this to be the case, one should observe that as a large number of firms and more innovative enterprises cluster in certain rural areas – thus attracting workers from other areas and pushing up population density – a larger number of routine entrepreneurs is 'pulled into' the market by the increased demand for

their goods and services. However, the increase in the share of replicative entrepreneurs as the population density increases should be smaller than the increase in the incidence of net firm creation or innovative enterprises, so that the differences between these proxies shrink.

Digging deeper, we find that this is not the case: while the incidence of net firm creation and innovation is not significantly related to population density in rural areas, the share of 'Real Entrepreneurs' is significantly and *negatively* associated with this variable, showing that the least populated rural areas have larger shares of self-employed workers. This finding reinforces our previous conclusion that a high incidence of self-employment in rural areas mostly captures last-resort choices dictated by lack of employment opportunities.[12]

6. Conclusions

Economists and policy makers alike consider entrepreneurs a crucial 'ingredient' in determining a country's or a region's economic prosperity. This is because entrepreneurs are thought to be conveyors of innovation, engines for job creation and sparks for technological change, economic growth and development. Unsurprisingly, a large empirical and theoretical literature on the characteristics and functions of the entrepreneurs, as well as on the effects of dense entrepreneurial environments, has emerged over the recent decades, and since the influential writings of Schumpeter (1921) and (1942).

Similarly, policy makers' interest in studying small business creation and designing interventions that stimulate entrepreneurial start-ups is always very high, and more so in the aftermath of the Great Recession. With the possibility of fiscal stimuli being progressively eroded by the need for a sustainable long-term path in public finances, the dynamics of private sector entrepreneurs are even more tightly associated with the prospects of different countries' swift recoveries or long lasting period of sluggish growth.

Despite the self-evident interest and importance of the figure of the entrepreneur in policy making and economics thinking, relatively little conclusive evidence has been gathered on the subject. This is because research in the field is hampered by the fundamental issue of defining and identifying who the entrepreneurs are. While the vast majority of the empirical investigations in this area rely on self-employment data to study entrepreneurship, the link between these two variables is far from proven.

In this paper, we have shed some light on this issue by looking at the correlation between the incidence of self-employment at the TTWA level and some of the most noticeable aspects of entrepreneurship, namely business creation and firms' innovative behaviour. To the best of our

[12] Note that the relation between population and self-employment, firm creation and innovation is always positive in urban areas, although only significant for the share of 'Real Entrepreneurs'.

knowledge, our study is the first to systematically investigate whether the common practice of using self-employment rates as a proxy for entrepreneurship has any validity using variation within-country and across economically relevant areas (i.e. TTWAs) in the incidence of self-employment, business start-up rates and innovation.

In order to carry out this analysis, we have first computed different measures of self-employment using UK labour force data. To being with, we created a crude measure of self-employment obtained by aggregating all self-employed individuals in a particular area. However, we also considered more refined ways of defining an entrepreneur by narrowing our definition to include only self-employed who: *(i)* occupy a managerial and professional position; *(ii)* own a business or a controlling majority of the business where they work; and *(iii)* are not freelancers or subcontractors. Our aim in doing this was to make sure that our results did not depend on the use of a specific definition of self-employment. We have then computed measures of business creation – both net and gross business start-up rates – as well as product and process innovation using census data on the universe of firms in Britain (BSD) and information contained in a representative survey of business innovative activities (CIS).

Our results show that there is a positive and significant correlation between the incidence of self-employment and business creation measured as both gross and net firm creation rates across urban TTWAs. Moreover, we find a positive and significant correlation between self-employment and innovation across urban areas. However, none of these results holds for rural TTWAs. Furthermore, our results do not appear to be driven by the sectoral composition of business activities across urban and rural areas. In fact, when we replicate our analysis separately for services and manufacturing across urban and rural TTWAs, we observe very similar patterns: a high incidence of self-employment in urban areas in either manufacturing or services is still positively correlated with higher net or gross firm creation, and either process (for services) or product (for manufacturing) innovation. On the other hand, these findings are not true for rural areas.

In order to explain the urban/rural dichotomy, we have exploited additional information contained in the LFS and constructed proxies for 'lack of employment opportunities'. Our concluding analysis shows that these measures significantly predict the misalignment between self-employment, business creation and innovation in rural areas, but not in urban areas. This pattern suggests that the urban/rural divide is related to lack of employment opportunities in rural areas which might push some workers into self-employment out-of-necessity. Previous studies (e.g. Alba-Ramirez, 1994 and Martinez-Granado, 2002) have documented that some self-employment spells can be explained by lack of alternative employment opportunities.

Nevertheless, to the best of our knowledge, no previous studies have linked this phenomenon to an urban/rural divide.

We believe our results carry important implications for the academic debate in the field which has widely used information on individuals' self-employment status to identify entrepreneurs. As long as the analysis focuses on urban labour markets – roughly speaking equivalent to the US Metropolitan Statistical Areas (MSAs) – our results suggest that self-employment could provide a relatively good proxy for entrepreneurship. However, our findings suggest that the same might not be true for rural TTWAs, and that an un-critical use of self-employment rates to approximate business creation or innovative activities could provide a distorted picture of the spatial distribution of entrepreneurship within a country or region.

Furthermore, our findings carry profound implications for public policies that promote self-employment with the aim of stimulating business creation and innovation. Indeed, this paper has documented an important urban/rural divide in individuals' motivation behind the decision to become self-employed. Our results therefore challenge the current governmental stance that tends to consider the incidence of self-employment in both urban and rural TTWAs as expressions of the same positive economic phenomenon.

References

Acs, Z. and D. Audretsch (2003): *Handbook of Entrepreneurship Research*, Kluwer Academic Publishers, London.

Alba-Ramirez, A. (1994): "Self-Employment in the Midst of Unemployment: The Case of Spain and the United States," *Applied Economics*, 26(3), 189–204.

Ardagna, S. and A. Lusardi (2008): "Explaining International Differences in Entrepreneurship: The Role of Individual Characteristics and Regulatory Constraints", in *International Differences in Entrepreneurship*, J. Lerner and A. Schoar, eds., Harvard Business School and MIT, Cambridge, USA.

Baumol, W. (2011): "Toward Prosperity and Growth," extract from *Kauffman Thoughtbook 2011*, accessible at www.kauffman.org/researchandpolicy on 16/02/2012.

Bertrand, M. and A. Schoar, (2006): "The role of family in family firms", *Journal of Economic Perspectives* 20, 73-96.

Bertrand, M., S. Johnson, K. Samphantharak and A. Schoar, (2008): "Mixing Family with Business: A Study of Thai Business Groups and the Families Behind Them", *Journal of Financial Economics*, 88(3), 466-98

Blanchflower, D. and A. Oswald (1998): "What Makes an Entrepreneur?," *Journal of Labor Economics*, 16(1), 26–60.

Blanchflower, D. and C. Shadforth (2007): "Entrepreneurship in the UK", *Foundations and Trends in Entrepreneurship*, 3(4), 1–108.

Bracke, P., C. Hilber and O. Silva (2012): "Homeownership and Entrepreneurship", mimeo, London School of Economics.

Cagetti, M. and M. De Nardi (2009): "Estate Taxation, Entrepreneurship, and Wealth", *American Economic Review*, 99(1), 85-111.

Carroll, R., D. Holtz-Eakin, M. Rider and H. Rosen (2000): "Income Taxes and Entrepreneurs' Use Of Labor", *Journal of Labor Economics*, 18(2), 324-351.

Dopeso-Fernandez, R. (2010): "Marshallian Agglomeration Economies and Entrepreneurship: the Spanish case", mimeo, Universitat Autonoma de Barcelona.

Ellison, G., Glaeser, E. and W. Kerr (2010): "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns", *American Economic Review*, 100(3), 1195-1213.

Evans, D. and L. Leighton (1989): "Some Empirical Aspects of Entrepreneurship", *American Economic Review*, 79 (3), 519–535.

Evans, D. and B. Jovanovic (1989): "An Estimated Model of Entrepreneurial Choice Under Liquidity Constraints," *Journal of Political Economy*, 87(4), 808–827.

Glaeser, E. (2009): "Entrepreneurship and the City", in *Entrepreneurship and Openness: Theory and Evidence*, D. Audretsch, R. Litan and R. Strom (eds.), Edward Elgar Publishing, Northampton.

Glaeser, E. and W. Kerr (2010): "Local Industrial Conditions and Entrepreneurship: How Much of the Spatial Distribution Can We Explain?," *Journal of Economics and Management Strategy*, 18(3), 623–663.

Gibbons, S., H. Overman, and P. Pelkonen (2010): "Wage disparities in Britain: People or place?", *Spatial Economics Research Centre*, DP0060, October 2010.

Holtz-Eakin, D., D. Joulfaian, and H. Rosen (1994a): "Entrepreneurial Decisions and Liquidity Constraints," *RAND Journal of Economics*, 25(2), 334–347.

Holtz-Eakin, D., D. Joulfaian, and H. Rosen (1994b): "Sticking it Out: Entrepreneurial Survival and Liquidity Constraints," *Journal of Political Economy*, 102(1), 53-75.

Hurst, E. and A. Lusardi (2004): "Liquidity Constraints, Household Wealth, and Entrepreneurship," *Journal of Political Economy*, 112(2), 319-347.

Hurst, E. and B. Pugsley (2010): "Non Pecuniary Benefits of Small Business Ownership", mimeo, University of Chicago.

Kanbur, S. (1979): "Of Risk Taking and the Personal Distribution of Income," Journal of Political Economy, 87(4), 769-97.

Kihlstrom, R. and J. Laffont (1979): "A General Equilibrium Entrepreneurial Theory of Firm Formation Based on Risk Aversion," Journal of Political Economy, 87(4), 719-48.

Knight, F. (1921): *Risk, Uncertainty and Profit*, Houghton Mifflin Co, Boston.

Lazear, E. (2004): "Balanced Skills and Entrepreneurship", *American Economic Review*, 58(2), 208-211.

Marshall, A. (1890): *Principles of Economics*, London: Macmillan and Co., Ltd. (Revised Edition ed., 1920).

Martinez-Granado, M (2002): "Self-Employment and Labour Market Transitions: A Multiple State Model", CEPR DP 3661.

Michelacci, C. and O. Silva (2007): "Why So Many Local Entrepreneurs?," *Review of Economics and Statistics*, 89(4), 615-633.

Office for National Statistics. *Business Structure Database, 1997-2008*: Secure Data Service Access [computer file]. Colchester, Essex: UK Data Archive [distributor], March 2011. SN: 6697.

Office for National Statistics. *Community Innovation Survey, 2001, 2005, 2007* and *2009*: Secure Data Service Access [computer file]. Colchester, Essex: UK Data Archive [distributor], March 2011. SN: 6699.

Office for National Statistics, *Quarterly UK Labour Force Survey, 1995-2009*: Secure Data Service Access [computer file]. Colchester, Essex: UK Data Archive [distributor], March 2011. SN: 6727.

Sanandaji, T. (2010): "Self-Employment Does Not Measure Entrepreneurship", mimeo, University of Chicago.

Schumpeter, J. (1921): *The Theory of Economic Development*, Harvard University Press, Cambridge.

Schumpeter, J. (1942): *Capitalism, Socialism, and Democracy*, Harper Publisher, New York.

Silva, O. (2007): "The Jack-of-All-Trades Entrepreneur: Innate Skill or Acquired Talent", *Economics Letters*, 97, 118-123.

Storey, D.J. (1994): *Understanding the Small Business Sector*, Thomson Learning, London.

Storey, D.J. (2006). "Entrepreneurship, Small and Medium Sized Enterprises and Public Policies", in Van der Horst, R., King-Kanauni, S. and Duffy, S. (eds), *Keystones of Entrepreneurial Knowledge*, Blackwells, Oxford.

Tables

Table 1: Geographical Distribution of Self-Employment and Entrepreneurship – Urban and Rural Areas

	Urban Areas			Rural Areas		
	Mean	Std.Dev.	C. of V.	Mean	Std.Dev.	C. of V.
All Self-Employed (LFS)						
All sectors	0.119	0.022	18.56%	0.134	0.033	24.28%
Manufacturing only	0.054	0.025	47.20%	0.070	0.036	52.41%
Services only	0.110	0.017	15.78%	0.122	0.027	22.30%
Self-Employed in Managerial Professions ('Managers'; LFS)						
All sectors	0.055	0.015	27.28%	0.063	0.018	28.74%
Manufacturing only	0.021	0.012	56.99%	0.026	0.016	60.77%
Services only	0.067	0.015	23.05%	0.077	0.021	27.09%
Self-Employed Owning/Controlling Business ('Owners'; LFS)						
All sectors	0.086	0.018	21.35%	0.097	0.024	25.13%
Manufacturing only	0.051	0.021	40.25%	0.062	0.030	48.93%
Services only	0.081	0.016	20.02%	0.091	0.023	25.17%
Self-Employed, No Freelance/Subcontractors ('Real Entrepreneurs'; LFS)						
All sectors	0.104	0.020	18.91%	0.120	0.028	23.75%
Manufacturing only	0.048	0.022	46.76%	0.062	0.034	55.89%
Services only	0.098	0.017	17.09%	0.111	0.023	21.00%
Net Firm Creation (as Share of Existing Firms; BSD)						
All sectors	0.005	0.005	89.40%	0.005	0.005	92.69%
Manufacturing only	-0.024	0.007	29.93%	-0.015	0.008	49.10%
Services only	0.007	0.005	71.07%	0.006	0.005	94.71%
Firm Entry (as Share of Existing Firms, BSD)						
All sectors	0.146	0.006	4.31%	0.140	0.009	6.12%
Manufacturing only	0.101	0.008	7.92%	0.102	0.009	8.95%
Services only	0.144	0.006	4.32%	0.137	0.008	6.26%
Share of Innovative Firms (Product and Process Innovation; CIS)						
All sectors	0.304	0.025	8.16%	0.295	0.035	12.01%
Manufacturing only	0.461	0.055	11.96%	0.462	0.079	17.21%
Services only	0.284	0.028	9.78%	0.268	0.034	12.83%

Note: Shares calculated using Labour Force Survey (LFS), Spring Quarters 1995-2009; Business Structure Database (BSD) for the years 1997 to 2008; and Community Innovation Survey (CIS) 2001, 2005, 2007 and 2009. The exact number of years used in LFS changes for alternative definitions and depending on data availability. Different definitions of self-employment explained in the body-text. Sectors excluded from calculations as follows: Agriculture, Hunting and Forestry; Fishing; Public Utilities (Electricity, Gas and Water); Public Administration and Defence; Private Households with Employees; and Extra-Territorial Organization and Bodies. There are 79 Urban and 79 Rural Travel To Work Areas (TTWAs; some TTWAs were regrouped following Gibbons et al., 2010).

Table 2: Sector Distribution of Self-Employment and Entrepreneurial Ventures – Urban and Rural Samples

	Entrepreneurship proxy:				
Industry:	All Self-Employed (LFS)	'Managers' (LFS)	'Owners' (LFS)	'Real Entrepreneurs' (LFS)	Share of Firms (BSD)
Individuals/Firms in Urban Areas Only					
Mining and Quarrying	0.13	0.19	0.17	0.13	0.11
Construction	23.46	4.12	19.81	21.36	9.82
High-Tech Manufacturing	2.83	2.45	4.36	2.69	3.60
Low-Tech Manufacturing	4.47	3.74	5.58	4.18	4.69
Transport/Warehouse/Communication	7.84	2.35	5.93	7.97	4.45
Wholesale Trade	5.18	5.40	6.91	5.54	9.20
Retail Trade	9.18	12.00	11.00	9.34	11.81
FIRE	3.78	5.62	4.43	3.97	7.51
Accommodation/Food Services	3.59	5.64	3.94	3.62	6.82
Entertainment Services	5.74	10.71	4.50	5.33	4.16
Professional and R&D Services	17.78	28.30	20.80	18.85	24.42
Health Care Services	7.81	9.83	5.67	8.21	5.42
General Services	8.23	9.67	6.90	8.80	7.98
Individuals/Firms in Rural Areas Only					
Mining and Quarrying	0.16	0.15	0.15	0.13	0.23
Construction	24.86	3.70	21.54	23.63	11.96
High-Tech Manufacturing	3.21	2.45	4.53	3.07	3.19
Low-Tech Manufacturing	5.42	4.16	6.23	4.92	5.04
Transport/Warehouse/Communication	6.05	2.44	5.28	6.13	5.01
Wholesale Trade	5.73	6.72	7.51	6.16	9.60
Retail Trade	10.51	14.39	12.03	10.51	13.20
FIRE	3.37	5.14	3.99	3.63	6.31
Accommodation/Food Services	5.84	10.57	6.34	5.50	9.22
Entertainment Services	4.46	8.48	3.93	4.55	3.75
Professional Services	14.45	22.31	15.84	15.20	18.33
Health Care Services	7.62	9.78	5.74	8.01	5.91
General Services	8.31	9.72	6.89	8.55	8.24

Note: Cells tabulate percentages of self-employed people (LFS) and firms (BSD) operating in one of the listed sectors. Sectors have been regrouped using 2-digit SIC Code as follows. Mining and Quarrying: codes 10-14; Construction: code 45; High-Tech Manufacturing: codes 22-24, 29-35; Low-Tech Manufacturing: codes 15-21, 25-28, 36-37; Transport/Warehouse/Communication: codes 60-64; Wholesale Trade: codes 50-51; Retail Trade: code 52; FIRE: codes 65-71; Accommodation/Food Services: code 55; Entertainment Services: code 92; Professional Services: codes 72-74; Health Care Services: codes 85, 90; General Services: codes 80, 91, 93. Total number of individual-data observations (LFS) as follows. Urban sample: approximately 536,000, except for 'Real Entrepreneurs' (approx. N=397,000) and 'Owners' (approx. N=313,000). Rural sample: approximately 159,000, except for 'Real Entrepreneurs' (approx. N=118,000) and 'Owners' (approx. N=93,000). This difference is due to these variables not being available for all years. Total number of BSD observations as follows. Urban sample: approximately 21 million (76.9% of firms). Rural sample: approximately 6 million (23.1% of firms).

Table 3: Reasons for Choosing Self-Employment – Urban and Rural Samples

Variable	Individuals in Urban Areas Only		Individuals in Rural Areas Only	
	First Reported Reason	Any Reported Reason	First Reported Reason	Any Reported Reason
All Self-Employed				
Non Pecuniary Reasons	35.66	39.36	35.56	39.07
Generate Income	7.39	13.52	7.48	13.96
Business Idea/New Product	11.24	17.98	12.31	18.86
Join Family Business	3.88	4.51	5.33	6.15
Lack Emp. Opportunities	10.90	12.83	11.64	13.50
Other Reasons	30.94	39.01	27.69	36.65
Self-Employed in Managerial Professions ('Managers')				
Non Pecuniary Reasons	36.47	40.09	35.16	38.69
Generate Income	5.84	12.03	5.56	11.52
Business Idea/New Product	12.42	20.16	13.48	21.14
Join Family Business	4.01	4.85	5.49	6.64
Lack Emp. Opportunities	9.14	10.76	10.16	12.40
Other Reasons	32.11	40.16	30.15	39.63
Self-Employed Owning/Controlling Business ('Owners')				
Non Pecuniary Reasons	38.77	42.76	37.10	40.73
Generate Income	7.52	14.54	7.35	14.05
Business Idea/New Product	12.46	20.15	13.32	20.55
Join Family Business	4.33	5.04	5.77	6.66
Lack Emp. Opportunities	11.06	13.01	12.23	14.21
Other Reasons	25.87	34.36	24.22	32.98
Self-Employed, No Freelance/Subcontractors ('Real Entrepreneurs')				
Non Pecuniary Reasons	37.03	40.73	36.62	40.10
Generate Income	7.14	13.59	7.15	13.81
Business Idea/New Product	11.78	18.79	12.92	19.70
Join Family Business	4.25	4.94	5.78	6.70
Lack Emp. Opportunities	11.10	13.03	11.92	13.96
Other Reasons	28.70	36.85	25.62	34.69

Note: Cells tabulate percentages of self-employed people reporting they chose self-employment for one of the listed reasons. Groups of reasons created as follows. Non Pecuniary Reasons include: 'Wanted independence/a change'; 'Better conditions of work'; 'Family commitments/wanted to work at home'. Generate Income includes: 'Wanted more money'. Business Idea/New Product includes: 'Opportunity arose - capital, space, and equipment available'; 'Saw the demand/market'. Lack of Employment Opportunities includes: 'No jobs available (locally)'; 'Made redundant'. Join Family Business includes 'Joined the family business'; Other Reasons include: 'Nature of the occupation'; 'Other'; 'No other reason given'. Information only available for LFS Spring Quarters 1999, 2000 and 2001.

Table 4: Explaining Differences Between Self-Employment Incidence and Other Measures of Entrepreneurship

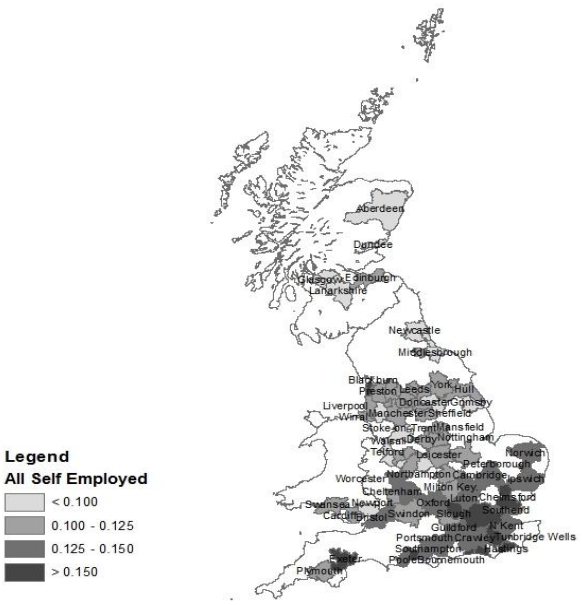
	Dependent Variable is the difference between:			
	(1)	(2)	(3)	(4)
	Self-Employment Incidence - Net Firm Creation	Self-Employment Incidence - Net Firm Creation	Self-Employment Incidence - Share of Innovative Firms	Self-Employment Incidence - Share of Innovative Firms
Panel A: 'Real Entrepreneurs' in Urban Areas				
Incidence of Underemployment	0.077	0.369	0.045	0.332
	(0.144)	(0.182)*	(0.151)	(0.200)
Incidence of Temporary Employment	-0.048	-0.116	-0.058	-0.101
	(0.102)	(0.195)	(0.148)	(0.197)
Lack of Full Time Employment	-0.033	0.206	-0.129	0.364
	(0.112)	(0.240)	(0.135)	(0.212)
Population density	0.097	0.254	0.103	0.224
	(0.079)	(0.153)	(0.080)	(0.166)
Panel B: 'Real Entrepreneurs' in Rural Areas				
Incidence of Underemployment	0.405	0.683	0.149	0.326
	(0.168)*	(0.179)*	(0.150)	(0.179)*
Incidence of Temporary Employment	0.272	0.337	0.299	0.242
	(0.179)	(0.195)*	(0.153)*	(0.197)
Lack of Full Time Employment	0.610	0.691	0.244	0.419
	(0.140)*	(0.206)*	(0.129)*	(0.195)*
Population density	-0.495	-0.491	-0.212	-0.354
	(0.162)*	(0.174)*	(0.198)	(0.182)*
Controls	No	Yes	No	Yes

Note: Regressions at the Travel to Work Area (TTWA) level. Number of observations: 79 in both urban and rural areas. Table reports coefficients from regressions of the dependent variable on the explanatory factors and robust standard errors in round parenthesis. *: 10% significant or better. Each cell corresponds to a different regression. Explanatory factors included one at the time. Descriptive statistics for the explanatory factors provided in Appendix Table 2. Controls include: unemployment rate; inactivity rate; share of adults with higher education; share of adults with A-level education (or equivalent); share of adults with O-level education (or equivalent); share of adults with other education (regressions using 'Lack of Full-Time Employment' further control for the incidence of part-time jobs). The dependent variables are constructed as the difference between the incidence of self-employment ('Real Entrepreneurs') standardized within the sample of urban areas and rural areas separately, and either net firm creation or share of innovative firms standardized within the same sample. The explanatory factors have also been standardized.

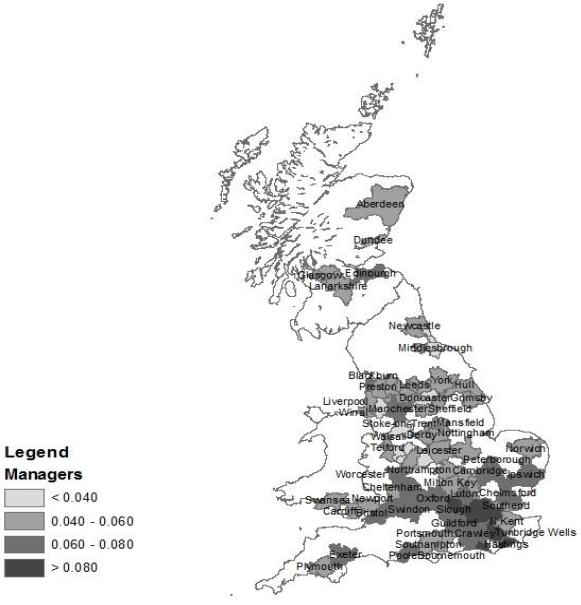
Figures

Figure 1.a: The Spatial Distribution of Self-Employment – Urban Areas

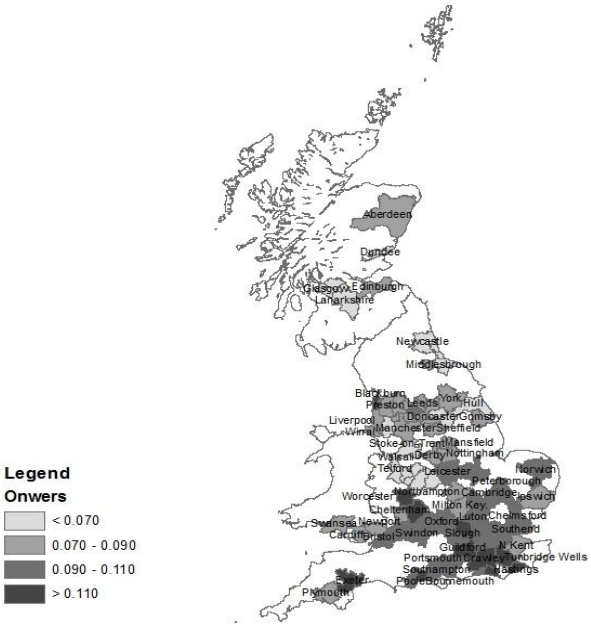
All Self-Employed



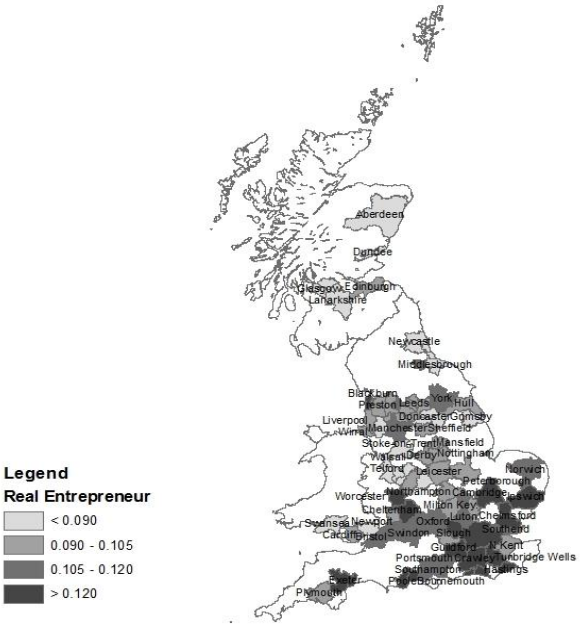
Self-Emp. in Managerial Professions ('Managers')



Self-Emp. Owning/Controlling Business ('Owners')



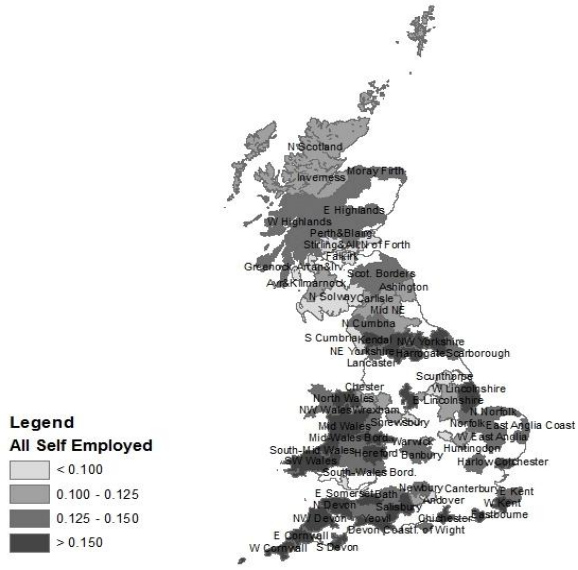
Self-Emp., No Freelance/Subcontractors ('Real Entrepreneurs')



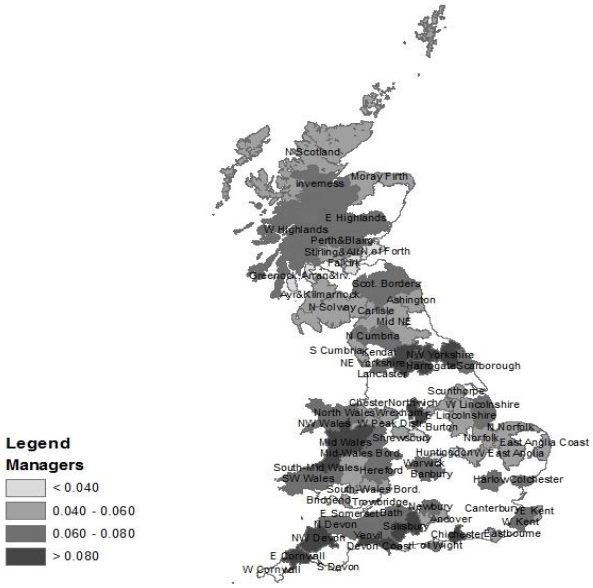
Note: Shares calculated using Labour Forces Survey, Spring Quarters 1995-2009. The exact number of years used changes for alternative definitions and depending on data availability. Different definitions of self-employment explained in the body-text. Sectors excluded from calculations as follows: Agriculture, Hunting and Forestry; Fishing; Public Utilities (Electricity, Gas and Water); Public Administration and Defence; Private Households with Employees; and Extra-Territorial Organization and Bodies. Urban Areas include to 79 Urban Travel To Work Areas (TTWAs; some TTWAs were regrouped following Gibbons et al., 2010; further: Inner and Outer London have been separated).

Figure 1.b: The Spatial Distribution of Self-Employment – Rural Areas

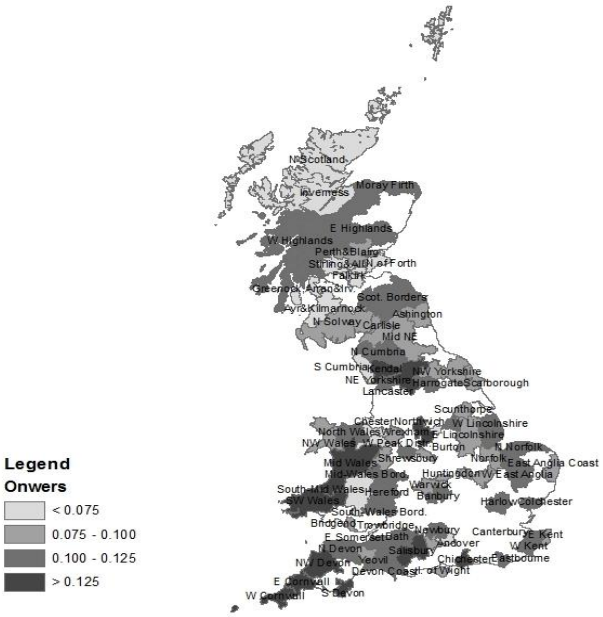
All Self-Employed



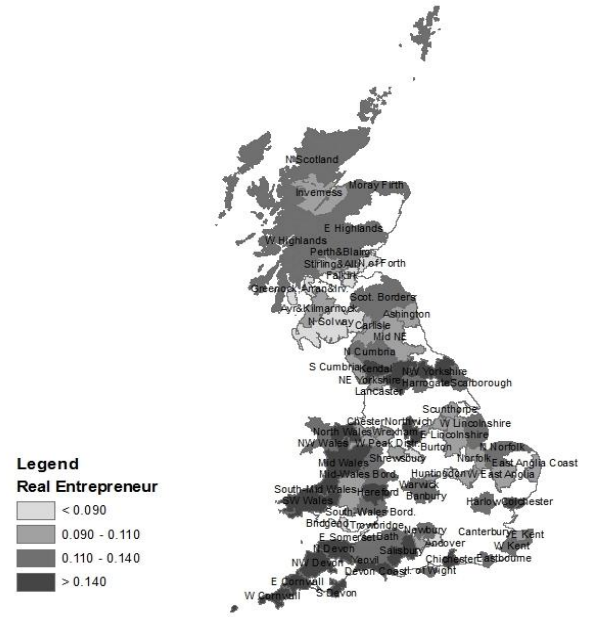
Self-Emp. in Managerial Professions ('Managers')



Self-Emp. Owning/Controlling Business ('Owners')



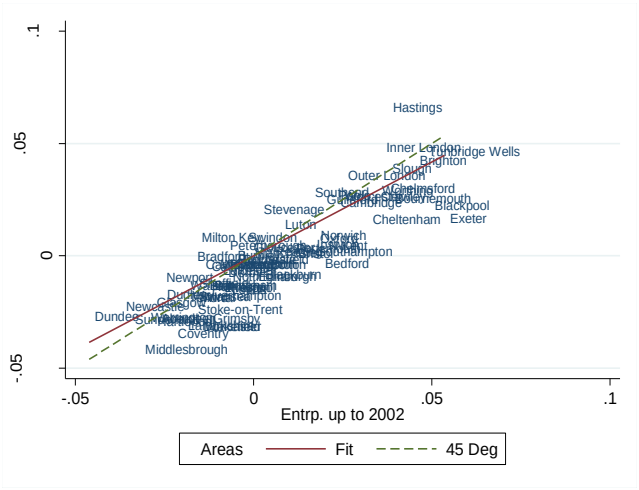
Self-Emp., No Freelance/Subcontractors ('Real Entrepreneurs')



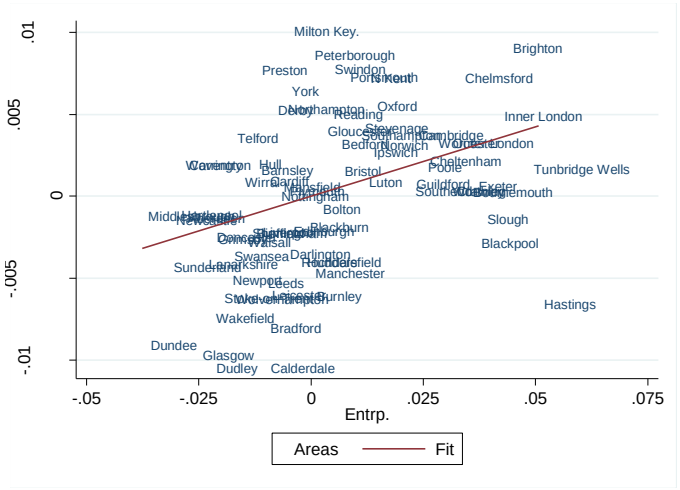
Note: Shares calculated using Labour Forces Survey, Spring Quarters 1995-2009. The exact number of years used changes for alternative definitions and depending on data availability. Different definitions of self-employment explained in the body-text. Sectors excluded from calculations as follows: Agriculture, Hunting and Forestry; Fishing; Public Utilities (Electricity, Gas and Water); Public Administration and Defence; Private Households with Employees; and Extra-Territorial Organization and Bodies. Urban Areas include to 79 Rural Travel To Work Areas (TTWAs; some TTWAs were regrouped following Gibbons et al., 2010).

Figure 2.a: All Self-Employed in Urban Areas

Time Alignment: up to and after 2002

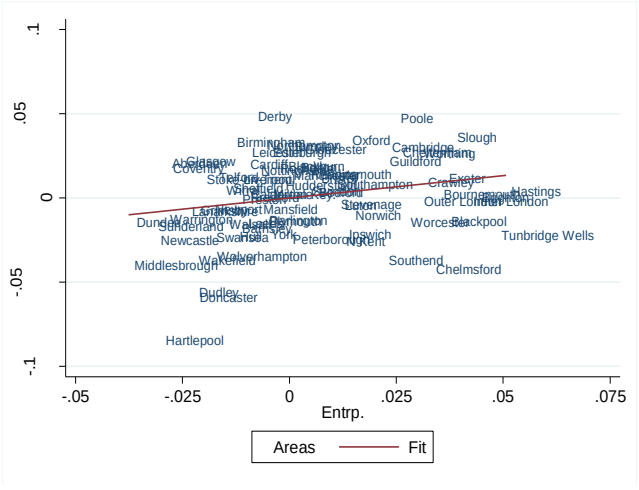


Self-Employment and Net Firm Creation



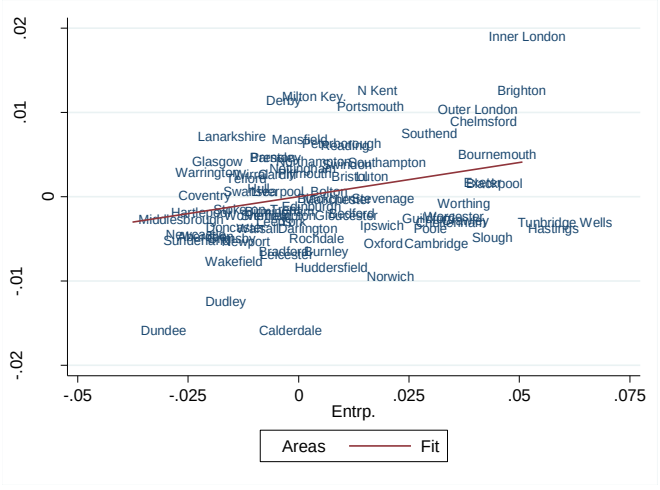
$y = 0.085\ (0.023)\ x + \varepsilon\ (R^2=0.152)$

Self-Employment and Innovation



$y = 0.266\ (0.134)\ x + \varepsilon\ (R^2=0.056)$

Self-Employment and Firm Entry

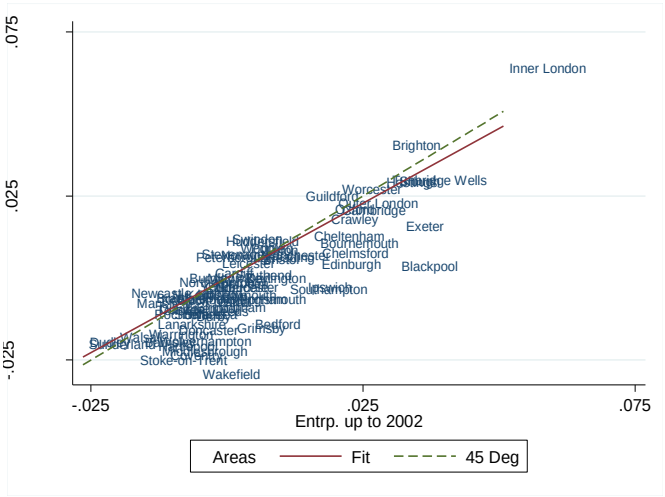


$y = 0.081\ (0.035)\ x + \varepsilon\ (R^2=0.080)$

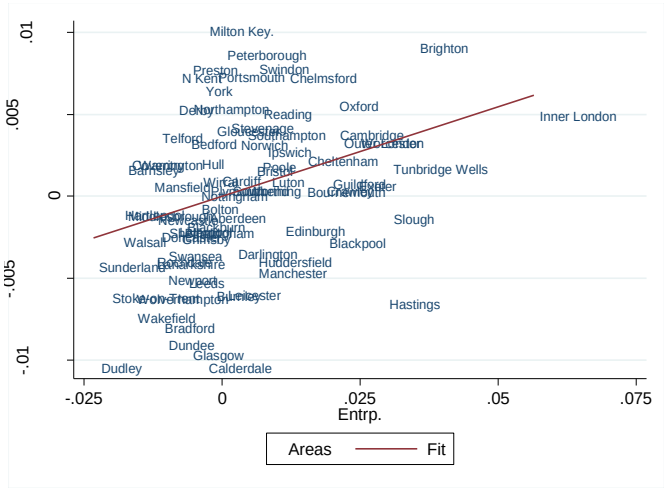
Note: Analysis includes 79 Urban TTWAs only. All variables displayed in the panels have been demeaned. Equations report regressions coefficient and standard errors of a regression of the variable on the vertical axis (*y*, e.g. share of innovative firms) on variable on the horizontal axis (*x*, i.e. a given measure of self-employment) plus a constant. Robust standard errors used.

Figure 2.b: Managerial Self-Employed ('Managers') in Urban Areas

Time Alignment: up to and after 2002

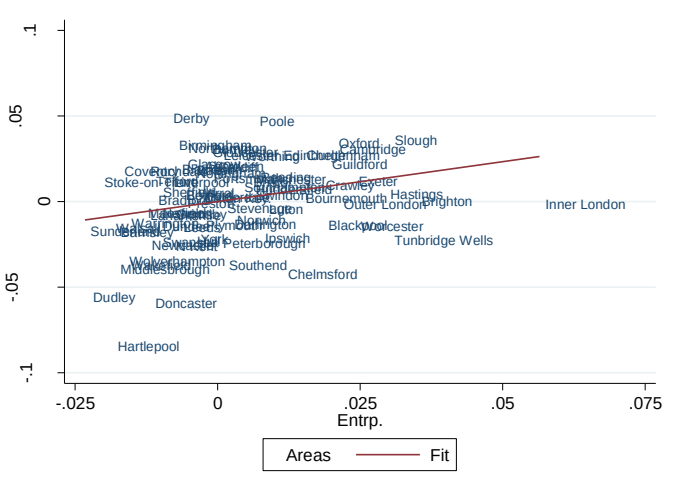


Self-Employment and Net Firm Creation

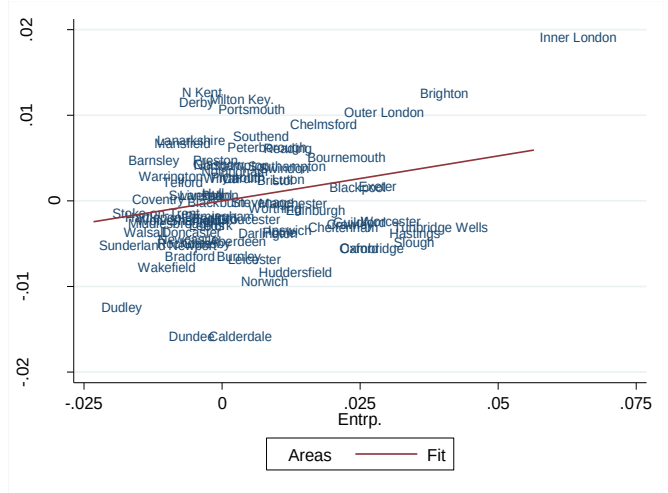


y =0.109 (0.030) x + ε (R^2=0.117)

Self-Employment and Innovation



Self-Employment and Firm Entry



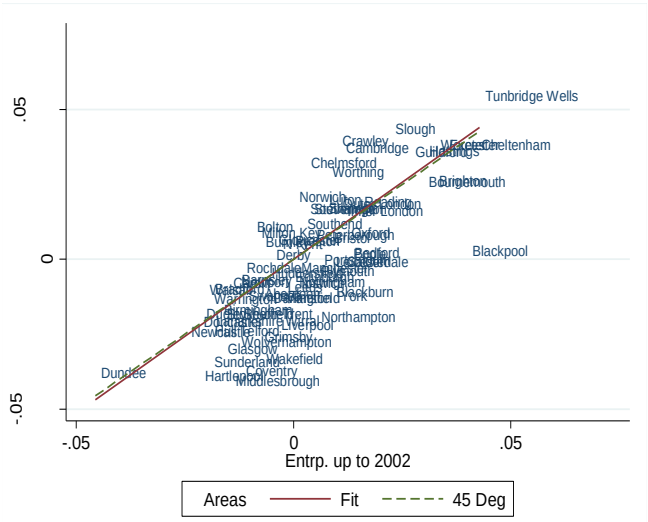
y = 0.464 (0.193) x + ε (R^2=0.080)

y = 0.105 (0.060) x + ε (R^2=0.063)

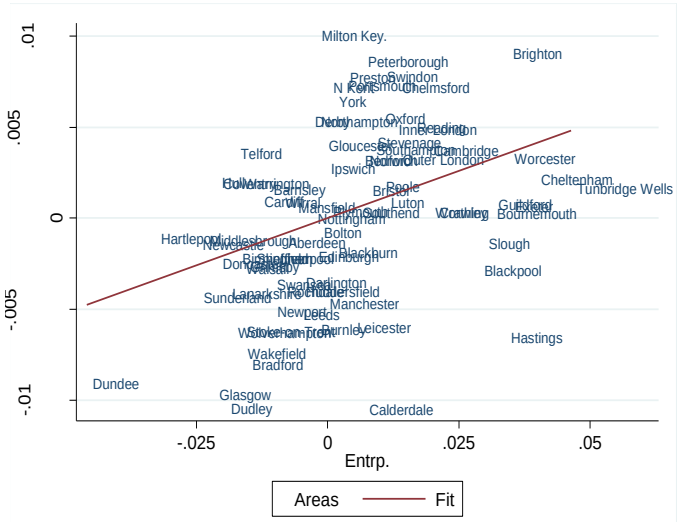
Note: Analysis includes 79 urban TTWAs only. All variables displayed in the panels have been demeaned. Equations report regressions coefficient and standard errors of a regression of the variable on the vertical axis (*y*, e.g. share of innovative firms) on variable on the horizontal axis (*x*, i.e. a given measure of self-employment) plus a constant. Robust standard errors used.

Figure 2.c: Self-Employed Owning/Controlling Business ('Owners') in Urban Areas

Time Alignment: up to and after 2002

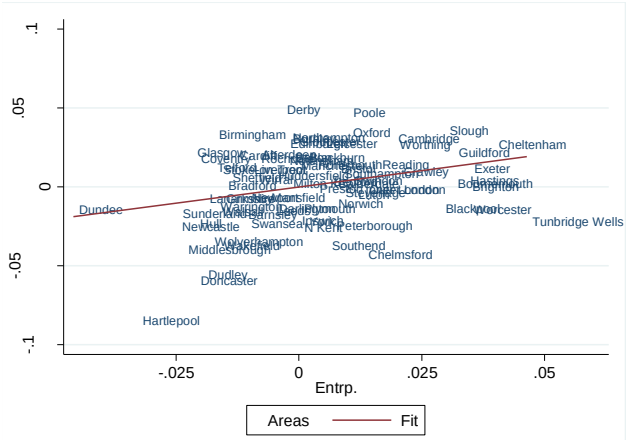


Self-Employment and Net Firm Creation



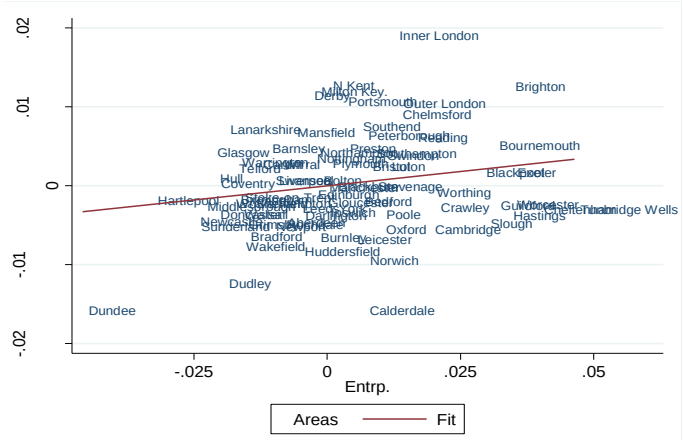
$y = 0.104\ (0.026)\ x + \varepsilon\ (R^2=0.156)$

Self-Employment and Innovation



$y = 0.415\ (0.165)\ x + \varepsilon\ (R^2=0.094)$

Self-Employment and Firm Entry

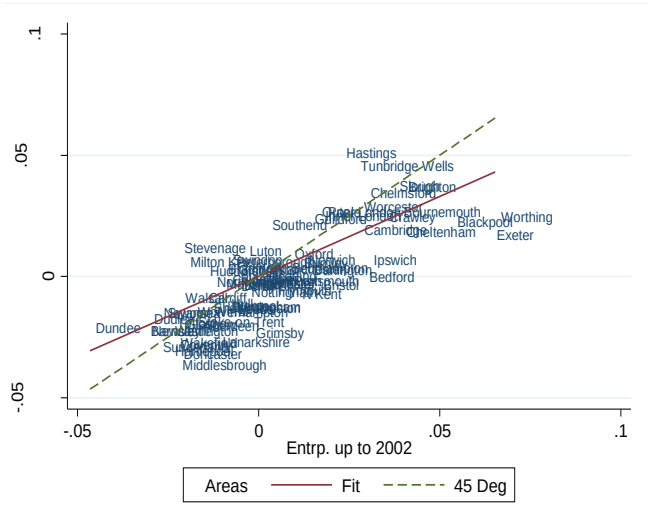


$y = 0.072\ (0.039)\ x + \varepsilon\ (R^2=0.044)$

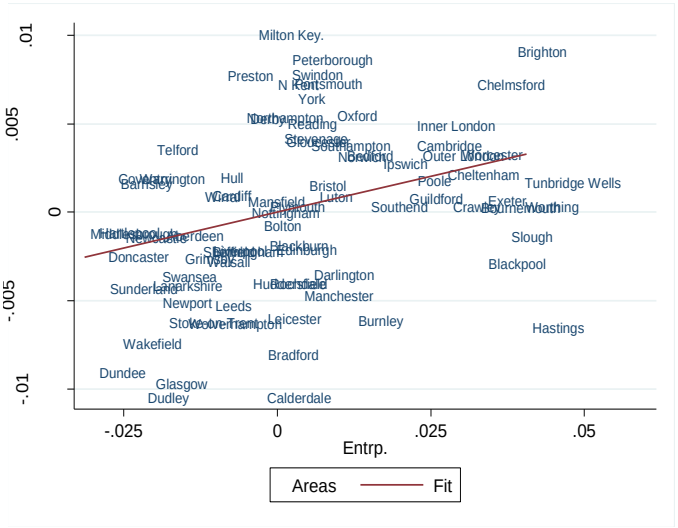
Note: Analysis includes 79 urban TTWAs only. All variables displayed in the panels have been demeaned. Equations report regressions coefficient and standard errors of a regression of the variable on the vertical axis (*y*, e.g. share of innovative firms) on variable on the horizontal axis (*x*, i.e. a given measure of self-employment) plus a constant. Robust standard errors used.

Figure 2.d: Self-Employed, No Freelance/Subcontractors ('Real Entrepreneurs') in Urban Areas

Time Alignment: up to and after 2002

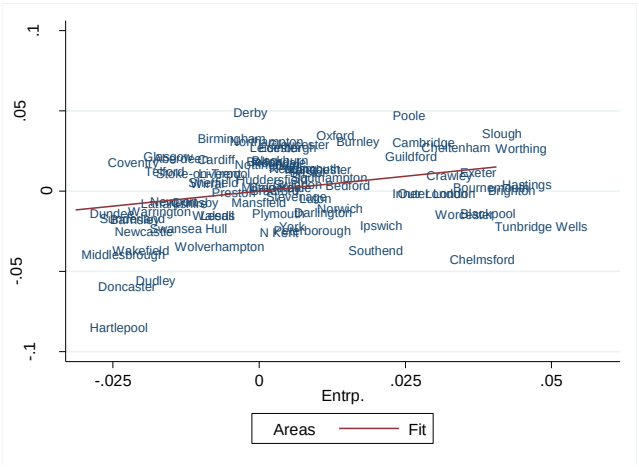


Self-Employment and Net Firm Creation



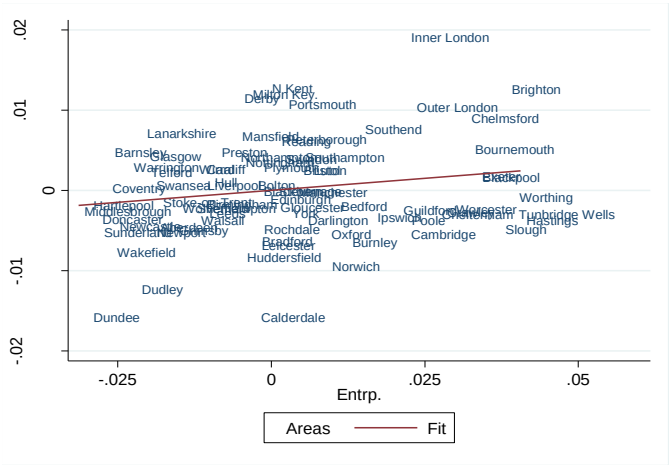
$y = 0.081\ (0.025)\ x + \varepsilon\ (R^2=0.110)$

Self-Employment and Innovation



$y = 0.373\ (0.155)\ x + \varepsilon\ (R^2=0.088)$

Self-Employment and Firm Entry

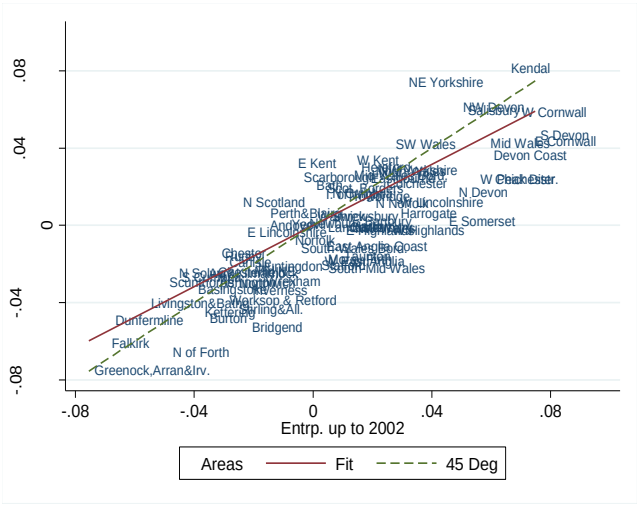


$y = 0.060\ (0.035)\ x + \varepsilon\ (R^2=0.035)$

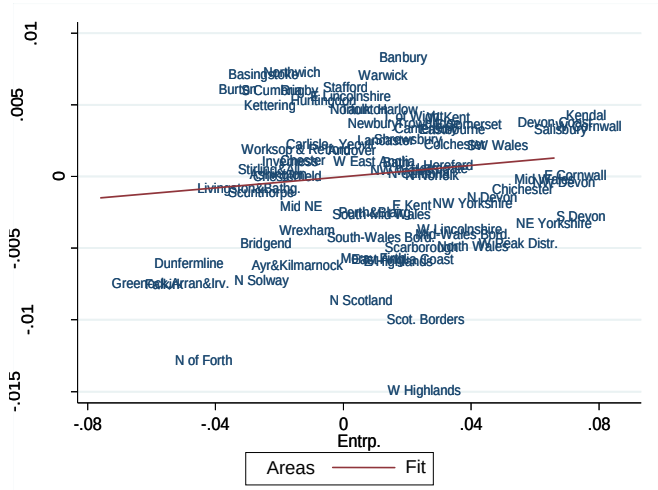
Note: Analysis includes 79 urban TTWAs only. All variables displayed in the panels have been demeaned. Equations report regressions coefficient and standard errors of a regression of the variable on the vertical axis (*y*, e.g. share of innovative firms) on variable on the horizontal axis (*x*, i.e. a given measure of self-employment) plus a constant. Robust standard errors used.

Figure 3.a: All Self-Employed in Rural Areas

Time Alignment: up to and after 2002

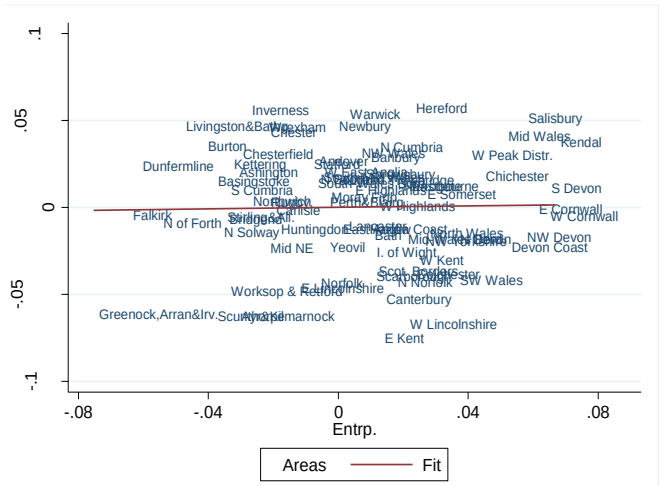


Self-Employment and Net Firm Creation



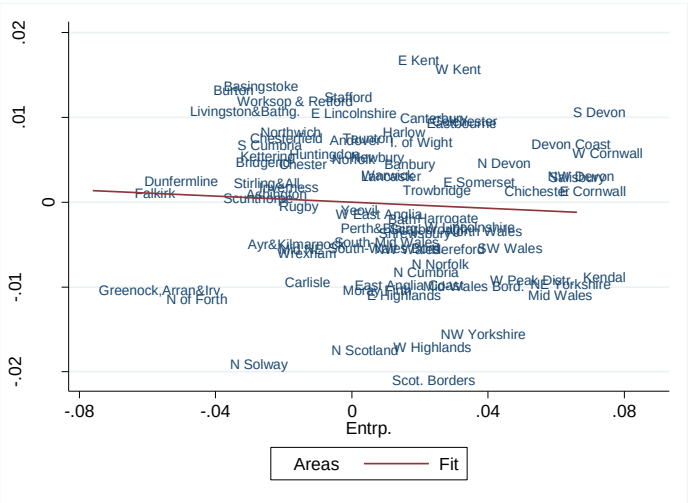
$y = 0.020\ (0.016)\ x + \varepsilon\ (R^2 = 0.017)$

Self-Employment and Innovation



$y = 0.022\ (0.117)\ x + \varepsilon\ (R^2 = 0.001)$

Self-Employment and Firm Entry

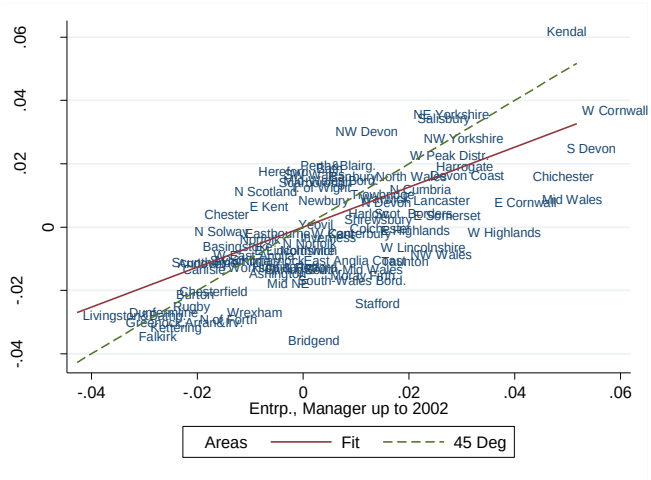


$y = -0.018\ (0.029)\ x + \varepsilon\ (R^2 = 0.005)$

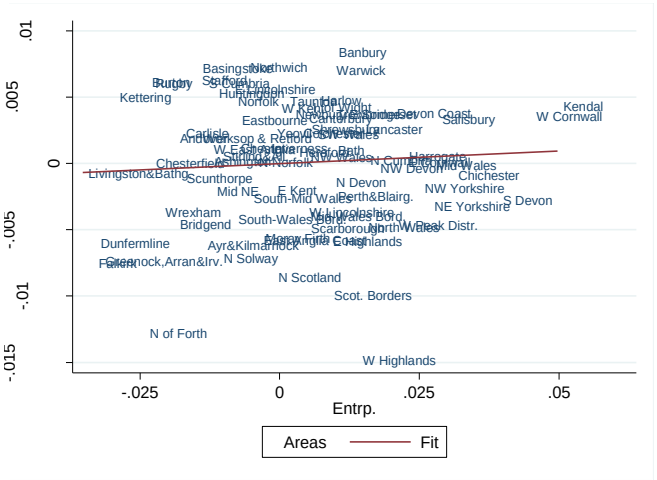
Note: Analysis includes 79 rural TTWAs only. All variables displayed in the panels have been demeaned. Equations report regressions coefficient and standard errors of a regression of the variable on the vertical axis (y, e.g. share of innovative firms) on variable on the horizontal axis (x, i.e. a given measure of self-employment) plus a constant. Robust standard errors used.

Figure 3.b: Managerial Self-Employed ('Managers') in Rural Areas

Time Alignment: up to and after 2002



Self-Employment and Net Firm Creation

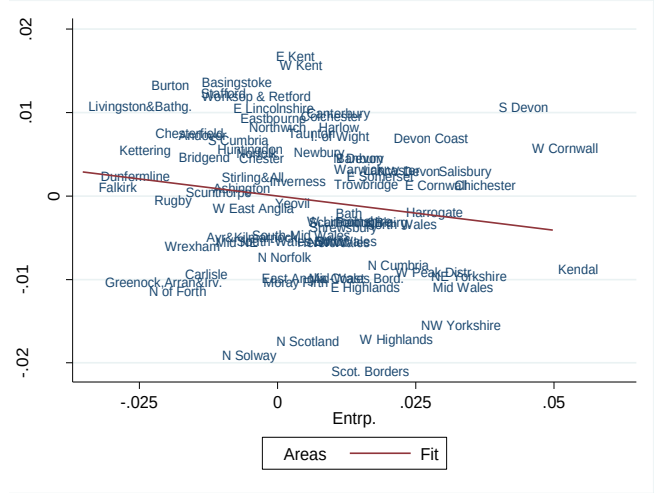


$y = 0.019\ (0.029)\ x + \varepsilon\ (R^2 = 0.005)$

Self-Employment and Innovation

$y = 0.130\ (0.193)\ x + \varepsilon\ (R^2 = 0.005)$

Self-Employment and Firm Entry

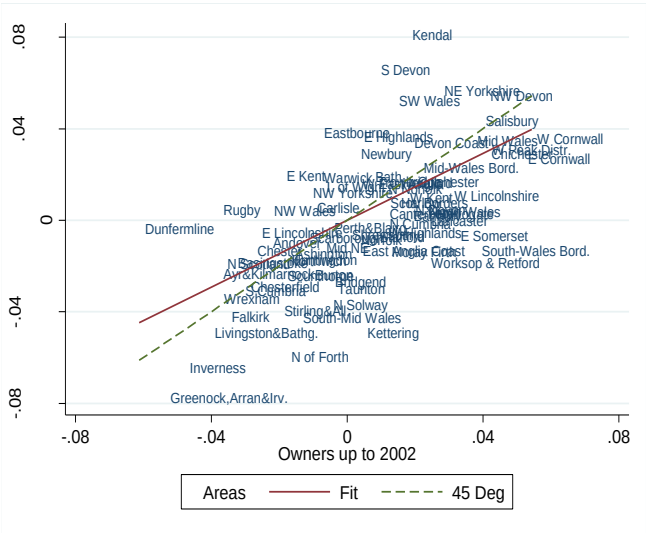


$y = -0.082\ (0.051)\ x + \varepsilon\ (R^2 = 0.030)$

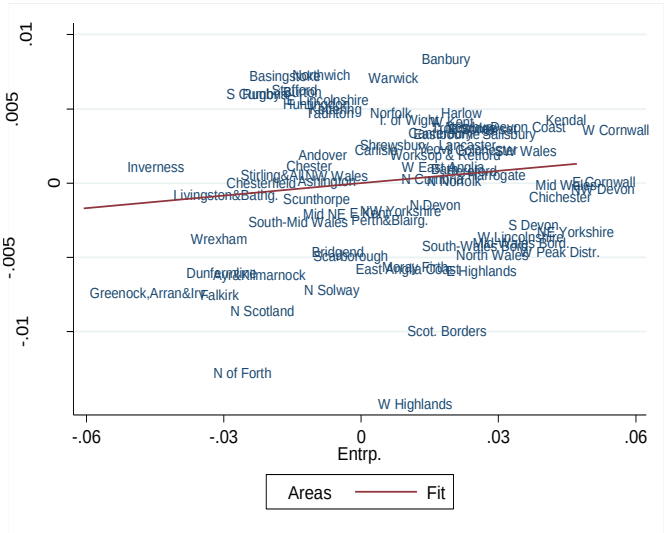
Note: Analysis includes 79 rural TTWAs only. All variables displayed in the panels have been demeaned. Equations report regressions coefficient and standard errors of a regression of the variable on the vertical axis (y, e.g. share of innovative firms) on variable on the horizontal axis (x, i.e. a given measure of self-employment) plus a constant. Robust standard errors used.

Figure 3.c: Self-Employed Owning/Controlling Business ('Owners') in Rural Areas

Time Alignment: up to and after 2002

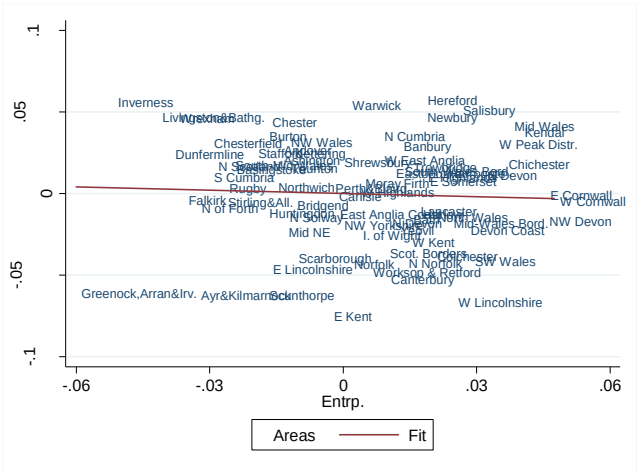


Self-Employment and Net Firm Creation



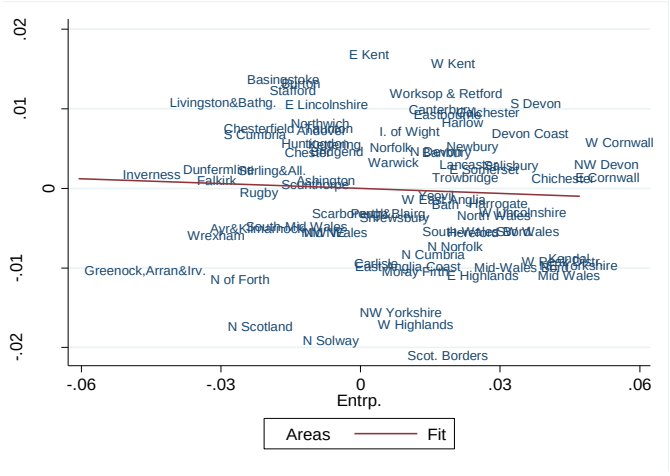
$y = 0.028\ (0.021)\ x + \varepsilon\ (R^2=0.020)$

Self-Employment and Innovation



$y = -0.068\ (0.172)\ x + \varepsilon\ (R^2=0.002)$

Self-Employment and Firm Entry

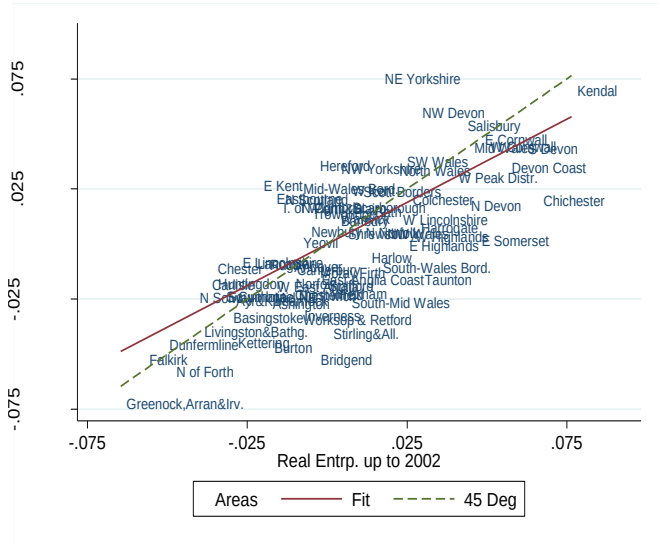


$y = -0.020\ (0.037)\ x + \varepsilon\ (R^2=0.003)$

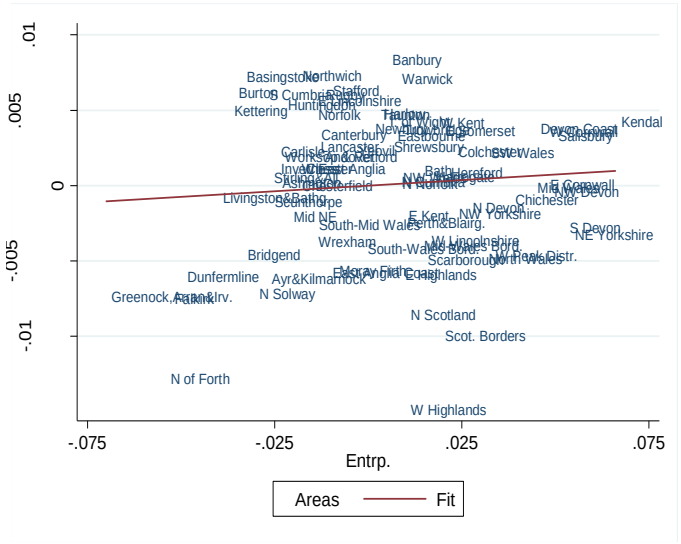
Note: Analysis includes 79 rural TTWAs only. All variables displayed in the panels have been demeaned. Equations report regressions coefficient and standard errors of a regression of the variable on the vertical axis (y, e.g. share of innovative firms) on variable on the horizontal axis (x, i.e. a given measure of self-employment) plus a constant. Robust standard errors used.

Figure 3.d: Self-Employed, No Freelance/Subcontractors ('Real Entrepreneurs') in Rural Areas

Time Alignment: up to and after 2002

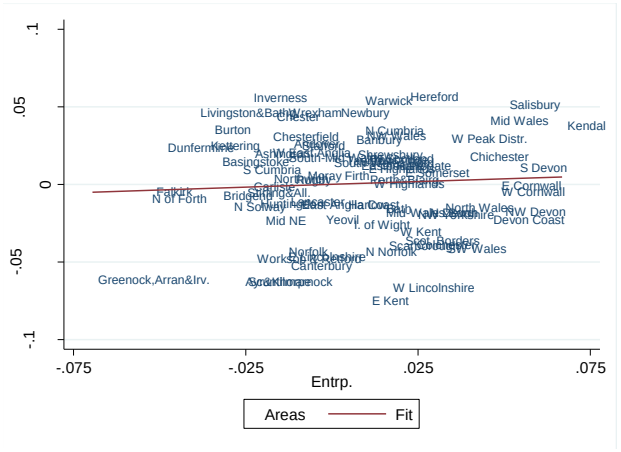


Self-Employment and Net Firm Creation



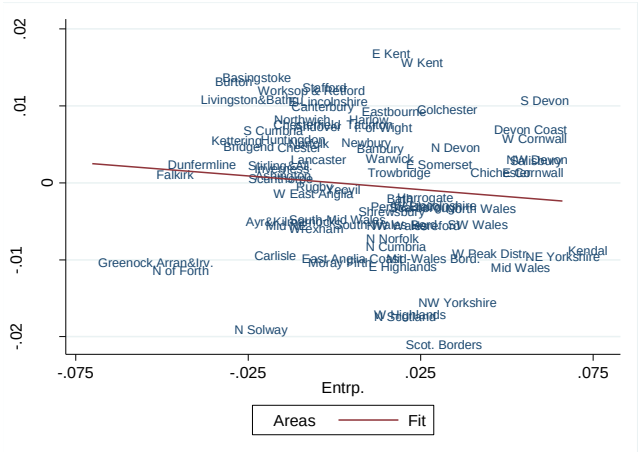
$y = 0.015\ (0.020)\ x + \varepsilon\ (R^2=0.008)$

Self-Employment and Innovation



$y = 0.071\ (0.133)\ x + \varepsilon\ (R^2=0.004)$

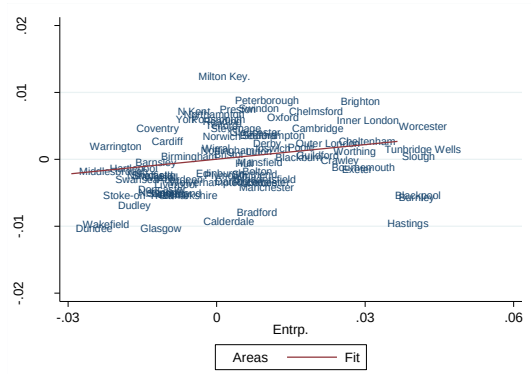
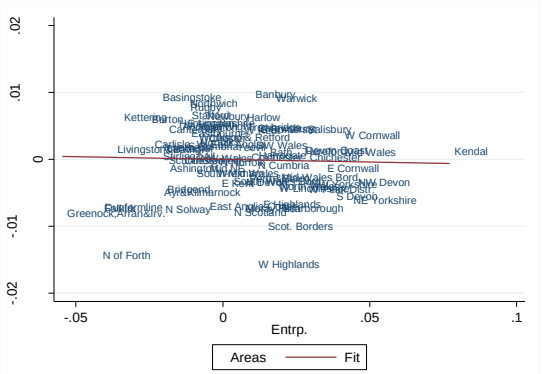
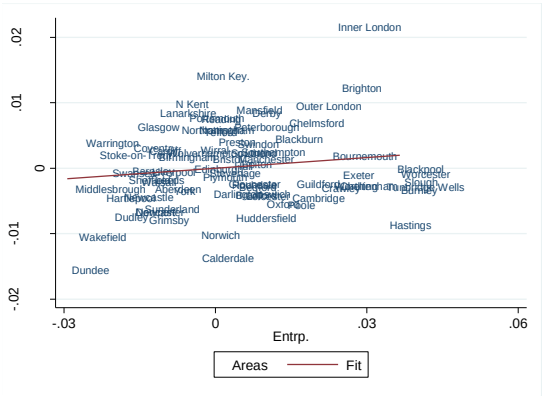
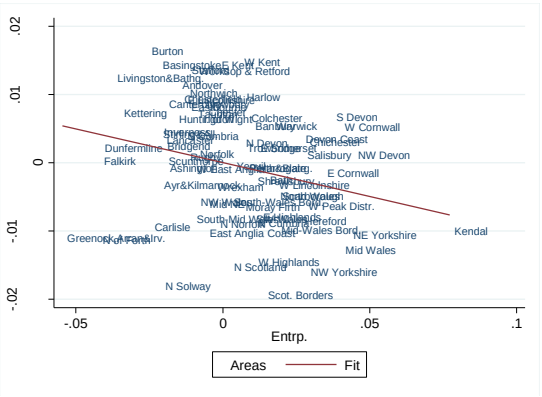
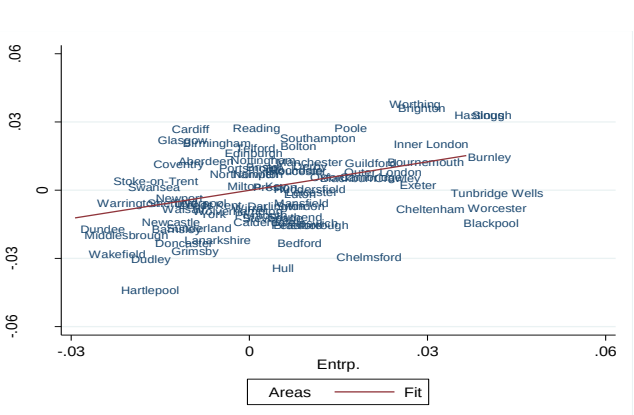
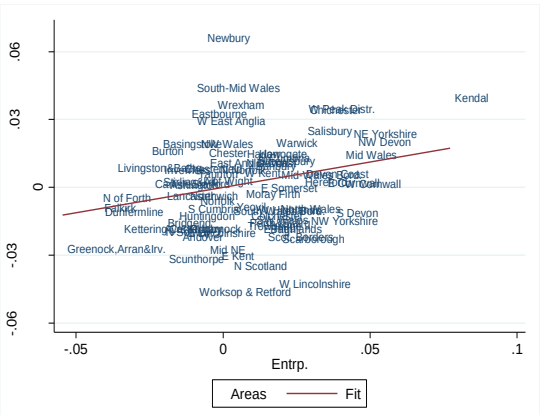
Self-Employment and Firm Entry



$y = -0.036\ (0.035)\ x + \varepsilon\ (R^2=0.014)$

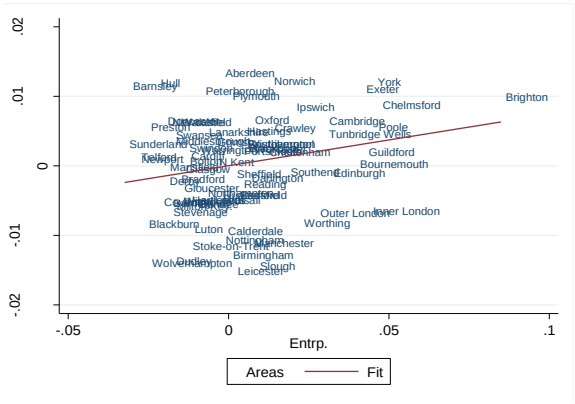
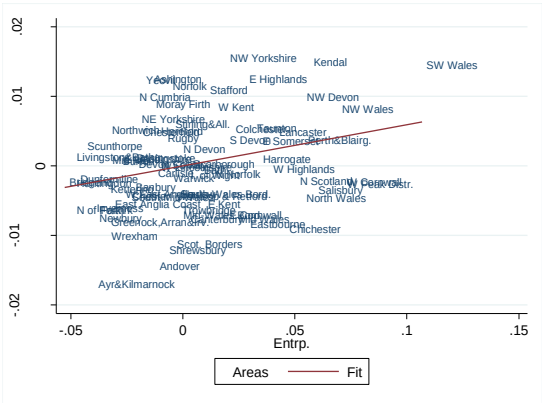
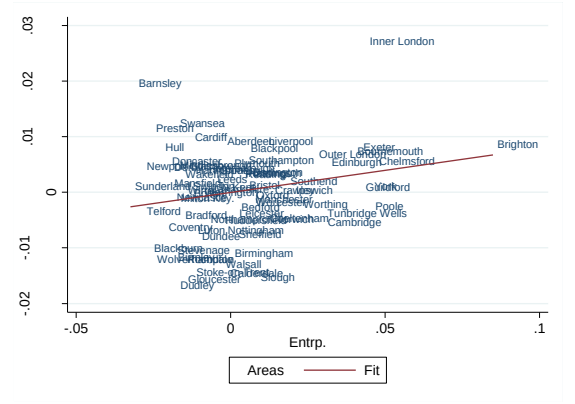
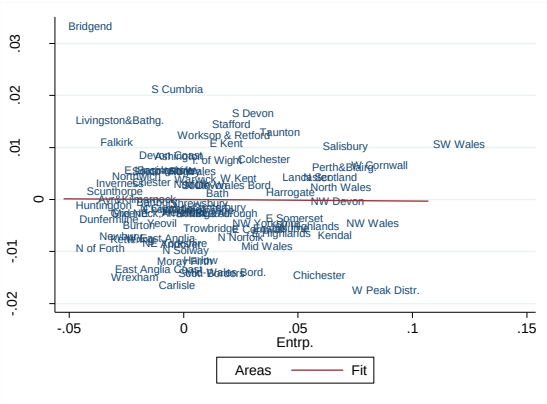
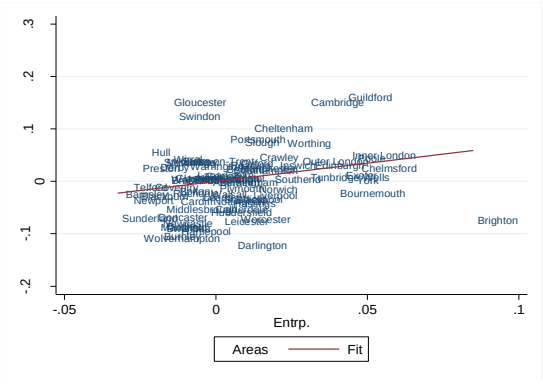
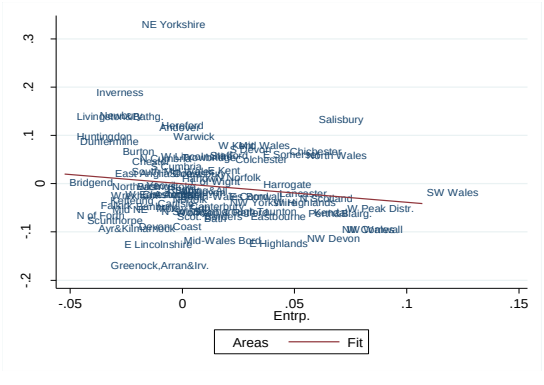
Note: Analysis includes 79 rural TTWAs only. All variables displayed in the panels have been demeaned. Equations report regressions coefficient and standard errors of a regression of the variable on the vertical axis (y, e.g. share of innovative firms) on variable on the horizontal axis (x, i.e. a given measure of self-employment) plus a constant. Robust standard errors used.

Figure 4: 'Real Entrepreneurs' Working in *Services* in Urban and Rural Areas

URBAN AREAS	RURAL AREAS
Self-Employment and Net Firm Creation	*Self-Employment and Net Firm Creation*
	
y = 0.073 (0.036) x + ε	*y = -0.008 (0.028) x + ε*
Self-Employment and Firm Entry	*Self-Employment and Firm Entry*
	
y = 0.055 (0.045) x + ε	*y = -0.099 (0.042) x + ε*
Self-Employment and Process Innovation	*Self-Employment and Process Innovation*
	
y = 0.414 (0.118) x + ε	*y = 0.224 (0.081) x + ε*

Note: See notes to Figures 2 and 3 (various panels). Descriptive Statistics for Process Innovation as follows. Urban areas: mean=0.163; std.dev.=0.025. Rural Areas: mean=0.151; std.dev.=0.074.

Figure 5: 'Real Entrepreneurs' Working in *Manufacturing* in Rural and Urban Areas

URBAN AREAS	RURAL AREAS
Self-Employment and Net Firm Creation	*Self-Employment and Net Firm Creation*
	
y =0.074 (0.032) x + ε	*y =0.055 (0.022) x + ε*
Self-Employment and Firm Entry	*Self-Employment and Firm Entry*
	
y =0.080 (0.042) x + ε	*y = -0.003 (0.035) x + ε*
Self-Employment and Product Innovation	*Self-Employment and Product Innovation*
	
y = 0.695 (0.386) x + ε	*y = -0.384 (0.258) x + ε*

Note: See notes to Figures 2 and 3 (various panels). Descriptive Statistics for Product Innovation as follows. Urban areas: mean=0.404; std.dev.=0.057. Rural Areas: mean=0.385; std.dev.=0.074.

Appendix Tables

Appendix Table 1: Entrepreneurs and Geographical Location –
Working in Same Local Authority and Working from Home/Using Home as Base – Urban and Rural Samples

	URBAN AREAS ONLY				RURAL AREAS ONLY			
	(1)	(2)	(3)	(4)	(5)	(6)	(9)	(8)
	Same LAD	Same LAD	Work/Use Home	Work/Use Home	Same LAD	Same LAD	Work/Use Home	Work/Use Home
Panel A: All Self-Employed								
Self-Employed (dummy 0/1)	0.227	0.273	0.507	0.474	0.169	0.201	0.541	0.501
	(0.013)	(0.015)	(0.005)	(0.005)	(0.007)	(0.007)	(0.005)	(0.005)
Panel B: Managers								
Self-Employed (dummy 0/1)	0.158	0.230	0.400	0.370	0.123	0.161	0.426	0.385
	(0.011)	(0.020)	(0.009)	(0.009)	(0.006)	(0.007)	(0.007)	(0.007)
Panel D: Owners								
Self-Employed (dummy 0/1)	0.189	0.240	0.443	0.392	0.146	0.181	0.479	0.419
	(0.013)	(0.013)	(0.005)	(0.005)	(0.007)	(0.007)	(0.006)	(0.006)
Panel C: Real Entrepreneurs								
Self-Employed (dummy 0/1)	0.238	0.275	0.508	0.467	0.181	0.209	0.544	0.497
	(0.013)	(0.013)	(0.006)	(0.005)	(0.008)	(0.008)	(0.006)	(0.005)
Year/Month Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demo + Job Controls	No	Yes	No	Yes	No	Yes	No	Yes
SIC 92 (2 digits) Effects	No	Yes	No	Yes	No	Yes	No	Yes
TTWA Effects	No	Yes	No	Yes	No	Yes	No	Yes

Note: See Appendix Table 2 for list of controls. Total number of observations as follows. Urban sample: approximately 536,000, except for 'Real Entrepreneurs' (approx. N=397,000) and 'Owners' (approx. N=313,000). Rural sample: approximately 159,000, except for 'Real Entrepreneurs' (approx. N=118,000) and 'Owners' (approx. N=93,000). Differences are due to variables not being available for all years. Standard errors clustered at the travel to work area (79 TTWAs for both urban and rural areas). All coefficients significant at 5% level or better. Share of people working and living in the same Local Authority/District: urban areas= 0.582 (std.dev.=0.493); rural areas=0.726 (std.dev.=0.446). Share of people working from home/using it as 'base': urban areas= 0.105 (std.dev.=0.307); rural areas=0.121 (std.dev.=0.326).

Appendix Table 2: Descriptive Statistics of Control Variables – Urban and Rural Sample

Variable	URBAN AREAS ONLY		RURAL AREAS ONLY	
Demographic Controls:	*Mean*	*Standard Deviation*	*Mean*	*Standard Deviation*
Female	0.477	0.012	0.490	0.018
Head of Household	0.525	0.013	0.513	0.015
Age	39.69	0.649	40.53	0.753
White	0.954	0.044	0.987	0.009
Status: single/cohabiting	0.305	0.032	0.273	0.019
Status: married	0.576	0.030	0.603	0.022
Status: separated	0.026	0.005	0.027	0.007
Status: divorced	0.080	0.010	0.084	0.013
Status: widowed	0.013	0.002	0.013	0.003
Skills: No Qual.	0.119	0.024	0.122	0.021
Skills: Other	0.237	0.027	0.242	0.029
Skills: O-Levels and Equiv.	0.274	0.033	0.277	0.030
Skills: A-Levels and Equiv.	0.183	0.026	0.194	0.036
Skills: Higher Education	0.187	0.056	0.165	0.038
N. of Children: 0	0.553	0.026	0.555	0.023
N. of Children: 1	0.192	0.017	0.190	0.018
N. of Children: 2	0.187	0.014	0.188	0.017
N. of Children: 3+	0.068	0.009	0.067	0.011
Household size: 1	0.136	0.023	0.126	0.017
Household size: 2	0.558	0.020	0.574	0.025
Household size: 3	0.195	0.017	0.199	0.018
Household size: 4+	0.111	0.014	0.100	0.015
			0.199	0.018
Job and Home Ownership Controls:				
Full Time	0.733	0.024	0.715	0.028
Second Job	0.043	0.008	0.052	0.013
Home Owners	0.821	0.045	0.812	0.033
Public Renter	0.095	0.033	0.098	0.036
Private Renter	0.083	0.032	0.090	0.030

Note: There are 79 urban TTWAs and 79 rural TTWAs. Number of underlying individual-level number of observations: approximately 536,000 (urban) and 159,000 (rural). Individual controls averaged at the TTWA level from individual data. Age controlled in regression analysis using shares of categorical variables constructed as follows: group 1 (16-25); group 2 (26-30)**;** group 3 (31-35); group 4 (36-40); group 5 (41-45); group 6 (46-50); group 7 (51-55); group 8 (56+).

Appendix Table 3: Descriptive Statistics of Explanatory Variables for Self-Employment Incidence – Urban and Rural Sample

Variable:	Mean	Std.Dev.	C. of V.
Urban Sample Only:			
Incidence of Underemployment	0.085	0.009	10.92%
Incidence of Temporary Employment	0.062	0.010	16.89%
Lack of Full Time Employment	0.100	0.028	28.07%
Population density	800.2	1073.7	134.18%
Rural Sample Only:			
Incidence of Underemployment	0.091	0.014	15.57%
Incidence of Temporary Employment	0.064	0.012	18.47%
Lack of Full Time Employment	0.117	0.034	28.71%
Population density	180.6	140.1	77.57%

Note: Data at the TTWA level and aggregated from LFS individual level data. There are 79 urban TTWAs and 79 rural TTWAs. Variables constructed as follows. Incidence of Underemployment refers to the fraction of workers not looking for a job who would like to work longer hours given the opportunity. Incidence of Temporary Employment refers to the fraction of workers whose job is not permanent (seasonal; fixed period, fixed task; agency temping; casual type; other reasons). Lack of Full Time Employment refers to the fraction of part-time workers who chose to work part-time because they could not find full-time employment. Population density is defined as number of people as recorded by the GB Census 2001 per square kilometres.

BIS Department for Business Innovation & Skills

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Llywodraeth Cynulliad Cymru
Welsh Assembly Government

Spatial Economics Research Centre (SERC)
London School of Economics
Houghton Street
London WC2A 2AE

Tel: 020 7852 3565
Fax: 020 7955 6848
Web: www.spatialeconomics.ac.uk

SERC is an independent research centre funded by the Economic and Social Research Council (ESRC), Department for Business Innovation and Skills (BIS) and the Welsh Assembly Government.