SAME DIFFERENCE? ETHNIC INVENTORS, DIVERSITY AND INNOVATION IN THE UK

MAX NATHAN

London School of Economics and Spatial Economics Research Centre
LSE, Houghton St, London, WC2A 2AE, UK. m.a.nathan@lse.ac.uk

Abstract [126 words]

Ethnic inventors play important roles in US innovation, especially in high-tech regions like Silicon Valley. Do 'ethnicity-innovation' channels exist elsewhere? This paper investigates the issues, using a new UK patents microdata panel and the novel ONOMAP name classification system. Ethnicity could influence innovation via 'star' migrants, co-ethnic network externalities, or production complementarities from diverse inventor communities. I develop rich descriptives and run multiple tests. UK ethnic inventors are spatially concentrated, as in the US, but have significantly different characteristics, reflecting UK-specific geographical and historical factors. I find positive effects of South Asian and Southern European co-ethnic group membership on individual patenting. The diversity of inventor communities also helps raise inventors' productivity. I find no hard evidence that ethnic inventors crowd out majority groups.

JEL Classification: J15, J24, J61, O3, R11, R23
Keywords: ethnic inventors, innovation, cultural diversity, cities

Acknowledgements

Thanks to Gianluca Tarasconi and Francesco Lissoni at Università Bocconi for KITES-PATSTAT patent data, to Pablo Mateos at UCL for ONOMAP coding, and to the ONS Virtual Microdata Lab for LFS datasets. Thanks for advice and support to Henry Overman, Ian Gordon, Maria Carvalho, Riccardo Crescenzi, Giulia Faggio, Ben Gidley, Simona Iammarino, Carlo Menon, Rosa Sanchis-Guarner-Herrero, Carlos Vargas-Silva and Vivek Wadhwa. Participants at the 2012 EUROLIO Geography of Innovation workshop, 2011 Urban Economics Association Conference, and seminars at COMPAS, SERC and the Regional Studies Association all provided helpful comments on previous versions. Any remaining errors and omissions are my own.

I am grateful to the ESRC and the Department of Communities and Local Government for research support. The views in this paper are my own, and do not necessarily represent those of the Department or the ESRC.

1. Introduction

At first glance, ethnicity, diversity and innovation do not seem closely linked. However, in recent years there has been growing policy and public interest in the role of 'ethnic inventors' in innovation (Kerr and Kerr, 2011, Leadbeater, 2008, Page, 2007, Legrain, 2006). These discussions and debates have largely drawn on recent experience in the United States.

The role of US ethnic inventors is striking. Since the 1980s minority communities, particularly those of South / East Asian origin, have played increasingly important roles in ideas generation in US science and technology sectors (Chellaraj et al., 2005, Stephan and Levin, 2001). US ethnic inventors – often migrants – are spatially concentrated at city-region level (Kerr, 2008a). In high-tech US clusters like Silicon Valley, 'ethnic entrepreneurs' help connect South Bay firms to global markets, and are responsible for 52% of the Bay Area's startups (Saxenian, 2006). There are positive links between migrant populations and US regional patenting (Hunt and Gauthier-Loiselle, 2010, Peri, 2007). Diasporic communities appear to play important roles in the diffusion of knowledge both across US cities, and between US regions and 'home' countries (Kerr, 2009, Kerr, 2008b).

By contrast, very little is known about the role of 'ethnic inventors' on innovation in European countries. The UK case is particularly interesting to explore, as Britain has become substantially more ethnically diverse in recent decades. In England and Wales, the number of people from non-white ethnic groups grew by 53% between 1991 and 2001; non-'White British' groups now stand at one in six of the population (Office of National Statistics, 2011). Immigration has been a main driver, with a number of 'new migrant communities' forming since the early 1990s, and the share of migrants in the working-age population almost doubling (Wadsworth, 2010, Kyambi, 2005). Just before the 2007 Great Recession, the UK's net inflow of migrants was around 198,000 people per year. The current Coalition Government has pledged to reverse these trends, and intends to reduce net immigration 'to the tens of thousands' by 2015 with a cap on non-EU migrants (Casciani, 2011).

This paper explores whether UK innovation has benefited from rising cultural diversity, as it has in the US. Specifically, I ask:

1) Do ethnic inventors or co-ethnic groups influence patenting rates in the UK?
2) Does the cultural diversity of inventor groups influence patenting rates?
3) Do urban environments affect ethnicity- or diversity-innovation effects?
4) What are the impacts, if any, on 'majority' inventors?

Changing demography could affect innovation in at least three complementary ways. First, migrants or minority individuals may be positively selected on the basis of skills or entrepreneurial behaviour, although this needs to be distinguished from other human capital endowments (Borjas, 1987). Second, by lowering transaction costs, co-ethnic groups can accelerate within-group ideas generation and transmission (Docquier and Rapoport, 2011). However, discrimination may constrain knowledge spillovers. Third, cultural diversity may improve ideas generation across all groups, if the benefits of a larger set of ideas, perspectives outweigh trust or communication difficulties between those groups (Berliant and Fujita, 2009, Page, 2007, Alesina and La Ferrara, 2004). Finally, these channels may be more pronounced in urban areas because of the spatial clustering of minority communities, agglomeration economies, or both.

I use a new 12-year panel of patents microdata to explore these issues, and identify ethnic inventors using the novel ONOMAP name classification system. This generates rich descriptive data showing that UK ethnic inventors have significantly different characteristics from their US counterparts, with substantially greater European-origin communities as well as groups reflecting the UK's colonial history and recent migration dynamics. While UK ethnic inventors are spatially clustered, as in the States, many high-patenting areas do not have diverse inventor communities.

I then estimate a knowledge production function linking inventors' patenting activity to individual, group and area-level characteristics. I am able to explore all four 'population-innovation' channels, using historic patenting information to control for inventor-level heterogeneity, alongside controls for group, area, industry and time factors. I find that ethnic inventor status has no significant effect on individual patenting rates. Conversely I find some positive effects for members of specific co-ethnic groups: Indian, South Asian and Southern European inventors. I also find small positive effects of inventor group diversity on individual patenting activity. Effects on majority inventors are less clear: increasing ethnic diversity has some negative links to majority groups' patenting activity individual level, but I find no effects of crowding out at area level. Urban location has relatively small effects on individual patenting after other individual and area-level factors are included. Results survive extensive robustness checks, and my identification strategy likely under-estimates true effects.

The paper makes several contributions to the field. It is one of very few studies exploring multiple ethnicity-innovation channels, at individual, group and area level. As far as I am aware, this is the first research of its kind in Europe. It also adds to the growing empirical literature on immigration, ethnicity and innovation, and to the emerging field of inventor microdata analysis (OECD, 2009). The paper is structured as follows. Section 2 sets out key concepts, theory and evidence. Section 3 introduces the data and identification strategy. Section 4 provides descriptive analysis. Section 5 outlines the model and estimation issues. Sections 6-8 give results, extensions and robustness checks. Section 9 concludes.

2. Definitions, frameworks and evidence

2.1 Definitions / key terms

'Innovation' and 'ethnicity' are fuzzy concepts that need careful definition. Innovation divides into invention, adoption and diffusion phases (Fagerberg, 2005). My chosen measure of innovation, patenting, is primarily an indicator of invention (OECD, 2009). Specifically, I look at shifts in individual patenting rates, or 'inventor productivity'.

Ethnic identity is a multifaceted concept with objective, subjective and dynamic elements (Aspinall, 2009). Quantitative measures of identity are partial: they focus on identity's visible, objective components, assuming away self-ascription and endogeneity issues (Ottaviano et al., 2007). Quantitative researchers working with ethnic identity therefore need a 'least-worst' proxy. I use the ONOMAP system to develop two such measures (see section 3). The first proxy is the ethnic group classifications prepared by the UK Office of National Statistics (ONS). The ONS measures attempt to combine different aspects of ethnic identity, but operate at a high level of generality and tend to focus on 'visible minorities' such as

Black and Asian communities (Mateos et al., 2007).[1] I use 'geographical origin' as a second proxy. Geographical origin can offer very fine-grained information, albeit on one dimension of identity. Name data conflates migrants and their descendents, so that origin operates as a measure of geographical 'roots'.[2]

2.2 Theoretical framework

Conventional theories of innovation have relatively little to say to about ethnicity or the composition of inventor communities. For example, Schumpeter (1962) focuses on the individual 'entrepreneurial function' inside and outside firms, as a source of developing new ideas; national 'innovation systems' approaches explore relationships between firms and public institutions, such as government agencies and universities (Freeman, 1987); while spatial approaches focus on clustering of innovative activity due to agglomeration-related externalities, particularly local knowledge spillovers (Jaffe et al., 1993, Audretsch and Feldman, 1996).

Endogenous growth theories help bridge demography to innovation. Shifts in the technology frontier help determine economic development, while human capital stocks and knowledge spillovers influence technological progress (Romer, 1990). However, access to knowledge is

---

[1] The full set of ONS 1991 groups is White, Black Caribbean, Black African, Indian, Pakistani, Bangladeshi, Chinese and Other. 'Other' is now the second-largest category in the UK population.

[2] Although not national identity: the vast majority of those born in the UK think of themselves as British MANNING, A. & ROY, S. 2007. Culture Clash or Culture Club? The Identity and Attitudes of Immigrants in Britain. *CEP Discussion Paper 790.* London: Centre for Economic Performance, London School of Economics. More broadly, ethnicity, nationality, sexuality and class are all elements in a broader sense of self (Fanshawe and Sriskandarajah, 2010).

likely to be uneven across locations, sectors and social groups (Agrawal et al., 2008). In turn, this suggests ways in which individual characteristics, networks, group composition or location might exert influences.

Individual selection

From an economic perspective, migration decisions reflect expected returns: potential migrants balance out gains from migration and costs of moving abroad (Borjas, 1987). This implies that migrants are 'pre-selected', with skill and entrepreneurialism both salient factors (Wadhwa et al., 2007). Migrant / minority status may thus positively predict patenting rates, over and above other human capital attributes. The challenge is to distinguish ethnicity from human capital endowments and (say) industry / area characteristics.

Co-ethnic networks

Co-ethnic social networks – such as diasporas or trans-national communities – may provide network externalities that accelerate ideas transmission (Docquier and Rapoport, 2011, Agrawal et al., 2008). Social networks offer their members higher social capital and levels of trust, lowering transaction costs and risk, and positively influencing innovative activity (Rodríguez-Pose and Storper, 2006, Kaiser et al., 2011). Matching and learning economies may be present within the group ('enclave' activity) and between different groups ('middleman minority' activities) (Bonacich, 1973).

In a closed economy, minority individuals are less likely to match ideas than those in the majority group, since there are fewer within-group possible matches. If members of majority group(s) discriminate against minority groups, or if minority-majority social connections are thin, this will limit ‘middleman minority’ activity (Zenou, 2011). Under globalisation, co-ethnic communities may be more influential. Increasing numbers of businesses in high-cost countries are looking to relocate research and development (R&D) activity into lower-cost countries (Mowery, 2001, Archibugi and Iammarino, 2002, Cantwell, 2005, Yeung, 2009). Diasporic communities with members in high-cost ‘host’ countries can help relocation to lower-cost ‘home’ countries (Kapur and McHale, 2005, Saxenian and Sabel, 2008). This raises both the size of the innovating co-ethnic community and the rate of information flow between its members, in both ‘home’ and ‘host’ locations (Kerr, 2008b, Kerr, 2009).

Diversity

‘Cultural distance’ between economic agents may also influence innovation rates. Specifically, individual inventors in a group may benefit from group-level diversity if this brings a richer mix of ideas and perspectives. Berliant and Fujita (2009) model a system of firm-level knowledge creation, showing that worker heterogeneity can accelerate ideas generation through individual-level production complementarities. ‘Cognitively diverse’ teams exploit a larger pool of ideas, suggesting that cultural mix is a good proxy for cognitive diversity (Hong and Page, 2004, Hong and Page, 2001). On the other hand, group-level cultural diversity may have a negative effect if it leads to lower trust and poor communication between individuals – for example, because of language barriers, misunderstandings, discriminatory attitudes or both. Spillovers (and co-operation) will be limited, leading to fewer, lower-quality solutions (Alesina and La Ferrara, 2004). Fujita and Weber (2003) argue that positive diversity effects will be most likely observed in research-based or ‘knowledge-intensive’ activities – such as those leading to patenting.

Urban areas

We might observe bigger co-ethnicity and diversity effects on innovative activity in cities because of population mix, agglomeration economies or both. Innovative activity, migrant and minority communities tend to be spatially clustered in urban areas. Kerr (2008a) finds that US ethnic inventors are spatially concentrated, largely in the biggest urban agglomerations. Urban areas may also have positive or negative 'amplifying' effects. For example, cultural diversity may help foster Jacobian knowledge spillovers across sectors (Jacobs, 1969). The more cosmopolitan the urban population, therefore, the greater the potential for hybridisation (Hall, 1998, Gordon et al., 2007). Conversely, members of minority communities may be physically isolated in particular urban neighbourhoods. Spatial segregation may limit the opportunity for knowledge spillovers and interaction with other groups (Zenou, 2011).

2.3 Evidence

For each of these channels, 'ethnicity' has a theoretically ambiguous effect on innovation. In practice, there is some empirical evidence of net positive effects. For instance, Indo- and Chinese-American communities make disproportionate contributions to science and engineering, in terms of workforce membership as well as Nobel Prize counts, elections to scientific academies and patent citations (Stephan and Levin, 2001). Foreign graduate students and skilled immigrants have a significant positive link on patent applications and grants (Chellaraj et al., 2005), while the immigrant contribution to US patenting rose from 7.3% to 24.6% 1998-2006 (Wadhwa et al., 2007). However, Hunt and Gauthier-Loiselle

(2008) suggest that once education and industry characteristics are controlled for, effects of individual migrant status disappear.

Several case studies suggest that diasporas are important influences on knowledge flows (Bresnahan and Gambardella, 2004, Saxenian, 2006, Docquier and Rapoport, 2011). Saxenian finds that 82% of Chinese and Indian immigrant scientists exchange technological information with colleagues in 'home' countries. Jaffe and Tratjenberg (1999) find that countries with a common language have larger R&D spillovers and international patent citation rates. Kerr (2008b) finds that co-ethnic communities in 'host' countries positively influence industrial performance in 'home' countries. Patenting growth in US cities is also faster for technologies that depend heavily on communities of immigrant inventors (Kerr, 2009). By contrast, Agrawal et al (2008, 2011) compare co-ethnic and co-location effects on patent citations, finding that physical location is up to four times more important.

A handful of recent studies link workforce diversity and innovation in knowledge-intensive environments. Parrotta et al (2011) find positive effects of workforce cognitive and cultural diversity on Danish firms' patenting rates. However Ozgen et al (2011) find weaker links between cultural diversity and product/process innovation in 'white collar' Dutch firms. Maré et al (2011) find no systematic links between workforce characteristics and innovation among businesses in New Zealand. More broadly, management studies find a small but significant workplace 'diversity advantage' on business performance (Landry and Wood, 2008).

At area level, Peri (2007) finds that US states' share of foreign-born PhDs is positively associated with levels of patenting. Hunt and Gauthier-Loiselle (2010) find that immigrant population shares raise state-level patenting, and that these effects are greater than individual-level effects. Kerr and Lincoln (2010) find positive effects of skilled migrants on both ‘ethnic’ and overall innovative activity at urban level. Ozgen et al (2010), studying EU NUTS2 regions, find positive connections between migration, immigrant diversity and regional patenting. Niebuhr (2006) finds a positive link between the diversity of German regions and regional innovation, especially for highly skilled employees.

By contrast, there is very little empirical evidence for the UK. Fairlie et al (2009) find some support for co-ethnicity effects on British-Indian business performance, although innovation is not considered. Studying London firms, Nathan and Lee (2011) report some evidence that migrant company founders are more entrepreneurial than average, and that both management and workforce diversity help raise product and process innovation. But Basu (2002, 2004) suggests considerable variation in levels of entrepreneurship across minority communities. Qualitative work by Nakhaie et al (2009) confirms that co-ethnicity effects both vary significantly across groups, and are shaped by wider socio-economic contexts.

3. Data

I have three main data sources. Patents information comes from the European Patent Office (EPO), which is made available through the OECD PATSTAT database. Raw patent data cannot typically be used at inventor level, because of common/misspelled names, or changes of address: I use data cleaned by the KITES team at Bocconi University, which allows robust identification of individual UK-resident inventors (Lissoni et al., 2006). The raw patents data covers the period 1978-2007, dated by priority year, and contains geocoded information on 141,267 unique British-resident inventors and 123,030 patents with at least one British-resident inventor.[3] Ethnicity information is then derived from inventor names using the ONOMAP name classification system (see below). Finally, I combine this individual-level information with area-level controls, assembled from UK Labour Force Survey held in the Office of National Statistics Virtual Microdata Lab.

3.1 Working with patents data

Innovation divides into invention, adoption and diffusion phases (Fagerberg, 2005). Patents are primarily an indicator of invention (OECD, 2009), which I use to look at shifts in individual patenting rates, or 'inventor productivity'. Patent data has several advantages: it provides detailed information on geography and patent owners, and is available for long time periods at relatively low cost. Not all inventions are patented, however, and patents may not record everyone involved in an invention. Patents also have variable coverage across industries (with a well-known bias towards manufacturing) and are sensitive to policy shocks (Li and Pai, 2010, OECD, 2009). I am able to deal with most of these issues through careful identification and estimation strategies.

I make several changes to the raw data for identification purposes. First, following Hall et al (2001), I truncate the dataset by three years to end in 2004.[4] Second, I group patent observations in four-year 'yeargroups'. Invention is a process, not an event, and inventors typically work on an invention for some time before filing a patent. Past and current factors

[3] 'Priority dates' represent the first date the patent application was filed anywhere in the world. The OECD recommends using priority years as the closest to the actual time of invention (OECD, 2009). The full dataset has 160,929 unique UK-resident inventors: 19,492 observations lack postcode information.

[4] There is typically a lag between applying for a patent and its being granted. This means that in a panel of patents, missing values typically appear in final periods.

together may influence year-on-year patenting, making it difficult to fit lags. Following Menon (2009), I use the mean citation lag of EPO patents to proxy the invention process.[5] Third, the main regressions use unweighted patent counts; area-level analysis uses weighted patent counts to avoid double-counting innovative activity (OECD, 2009). Fourth, I use a combination of technology field dummies and area-level industrial structure controls to control for structural biases in patenting activity across different industrial sectors. Finally, I restrict the sample to 1993-2004, reserving pre-1993 inventor activity to construct individual-level controls (see section 5).

3.2 Identifying ethnic inventors

I use the ONOMAP name classification system to generate ethnicity information for individual inventors, building on similar approaches in pioneering US studies by Kerr (2008a) and Agrawal et al (2008). Originally designed for mining health service patient data, ONOMAP classifies individuals according to most likely cultural, ethnic and linguistic ('CEL') characteristics identified from forenames, surnames and forename-surname combinations. ONOMAP exploits similarities and differences between name families – for example, 'John Smith' is more likely to be ethnically British than French (Lakha et al., 2011).

ONOMAP is built from a very large names database drawn from UK Electoral Registers plus a number of other contemporary and historical sources, covering 500,000 forenames and a million surnames across 28 countries (Mateos et al., 2007). These are then algorithmically grouped together, combining information on geographical area, religion, language and

[5] If patent B cites patent A, the 'citation lag' between the two is the time period between the filing of A and the filing of B: the lag offers a rough way to capture the relevant external conditions affecting patenting. The mean citation lag for EPO patents is four years (OECD, 2009), so I group patents into four-year periods.

language family. Separate classifications of surnames, forenames and surname-forename combinations are produced.[6] ONOMAP has the advantages of providing information at several levels of detail and across several dimensions of identity. By using surname *and* forename information, it is also able to deal with many names with multiple cultural origins; the historically fuzzy boundaries of many states (e.g. Germany and the Netherlands), and the alteration and/or adoption of names traditional to the UK. ONOMAP also matches 99% of inventor names and has been extensively robustness-tested against UK Census data and other public sources.

I use ONOMAP's information on geographical origin and ONS ethnic groups. ONS ethnic group information is based on the nine categories developed for the 1991 Census, and as set out in Section 2, is quite limited. Geographical origin information provides objective, finer-grained information on twelve zones across Europe, Africa, Asia and the Americas, and is my preferred measure of ethnicity.[7]

ONOMAP has some limitations. First, as it observes only objective characteristics of identity, it identifies *most likely* ethnicity. Second, it does not distinguish migrants from their

[6] Names information is drawn from 1998 and 2004 GB Electoral Registers, Northern Ireland Electoral Register 2003, Irish Electoral Register 2003, plus electoral data from Australia (2002), NZ (2002), United States (1997) and Canada (1996). Experian MOSAIC geo-demographic data and Experian Consumer Dynamics data are used to boost the sample. This produces 25360 surnames and 299797 first names. These are classified using a combination of triage, spatio-temporal analysis, geo-demographic analysis, text mining, 'name-to-ethnicity' techniques from population registers and researching international name frequencies. 'British names' are those originating in the British Isles (including Ireland) or arriving there before 1700 (Mateos et al, 2007).

[7] The full set of twelve zones is Africa, Americas, British Isles, Central Asia, Central Europe, East Asia, Eastern Europe, Middle East, Northern Europe, South Asia, Southern Europe and Rest of the World.

descendents, so that geographical origin is best interpreted as ‘roots’. Third, there are a small number of conflict cases (including the author’s own name) where ONOMAP assigns the most probable identity. Finally, some countries share a common language, so that many North American and Australasian-origin inventors are largely identified as British-origin inventors (or unclassified). However, these groups represent a very small share of the UK minority population, so that potential measurement error is small.[8]

4. Descriptive analysis

Tables 1-4 provide some initial descriptive analysis. Table 1 breaks down inventors by CEL subgroup, showing the 30 largest groups. We can see that while English, Welsh, Scottish and Celtic[9] inventors make up the bulk of the sample, other inventor groups divide fairly evenly into geographically proximate communities (e.g. Irish, plus a series of European groups), groups reflecting the UK’s colonial history in South and East Asia (e.g. Indian Hindi, Sikh, Pakistani, Hong Kong Chinese) plus some largely recent migrant communities (e.g. Polish, Vietnamese).

Insert Table 1 about here

Table 2 recuts the sample by probable geographical origin zones and by 1991 ONS ethnic groups. Geographical origin zones (top panel) allow me to preserve some of the detail from

[8] LFS figures suggest working-age individuals of American, Canadian, Australian and New Zealand origin comprise 8.84% of UK migrants in 1994, 7.98% in 2004.

[9] ‘Celtic’ denotes names common to Scottish, Welsh and Irish CEL types.

the full CEL classification, including several areas of Europe as well as South and East Asia. As highlighted in the previous section, ONS ethnic groups (bottom panel) are much less flexible, focusing on visible majorities and minorities, relegating the rest of the inventors to 'other'.

Insert Table 2 about here

Table 3 presents location quotients for the 40 Travel to Work Areas with the largest shares of ethnic inventors by geographical origin, and confirms that ethnic inventors are more spatially clustered than the wider migrant population. High-ranking TTWAs are predominantly urban, although a number of rural areas also feature, predominantly university towns (St Andrews, Lancaster, Canterbury) or areas adjoining TTWAs with universities (Bude and Holsworthy) and/or manufacturing clusters (Holyhead, Pembroke and Tenby, Louth and Horncastle).[10]

Insert Table 3 about here

Table 4 gives weighted counts for the 40 TTWAs with the highest patenting activity: to minimise double counting, I weight each patent by the number of inventors involved. The results follow the familiar geography of UK innovative activity. A number of these high-patenting areas also have large ethnic inventor shares and diverse inventor groups (for

[10] Many inventors will work in professional / technical occupations, which are characterised by longer-than-average commuting distances. Building commuting zones on the basis of these workers' commuting patterns substantially reduces the total number of zones ROBSON, B., BARR, R., LYMPEROPOULOU, K., REES, J. & COOMBES, M. 2006. A framework for City-Regions: Working Paper 1: Mapping City-Regions. London: Office of the Deputy Prime Minister., suggesting that commuting across conventional TTWAs is not uncommon.

example London, Southampton, Crawley, Oxford and Cambridge). However, another group of high-patenting TTWAs have rather more homogenous inventor and general populations (for example, Bristol, Manchester, Reading and Ipswich).

Insert Table 4 about here

A number of broad lessons emerge from the descriptives. First, the UK's population of ethnic inventors appears substantially different from that of the US. American ethnic inventor communities are dominated by South and East Asian groups (Kerr, 2008a). By contrast, the UK has a number of European groups, with South Asian and East Asian inventors drawn in large part from former colonies. Second, as in the US ethnic inventors are spatially concentrated, and more clustered than minority populations in general. Third, not all high-patenting locations have large ethnic inventor shares or diverse inventor communities.

5. Model and estimation strategy

5.1 Identification

For the regression analysis I build a panel of UK-resident inventors' patenting activity between 1993 and 2004 inclusive – reserving the pre-1993 period for inventor-level human capital controls (see below). Each inventor-yeargroup cell records how many times an inventor patents in that time period. The basic panel covers 125,502 inventors across three four-year yeargroups, giving 376,506 observations. Cell counts vary from zero to 36, with a mean of 0.318.

I use postcode information to locate inventors in UK Travel to Work Areas (TTWAs), which are good approximations of local economies (and superior to administrative units such as local authority districts).[11] I then fit an urban / rural typology of TTWAs developed in Gibbons et al (2011), allowing me to explore the potential effects of urban environments.

Working with inventors in a panel data setting presents three linked identification challenges. First, inventors are only observed when patenting. Blanking all cells where the inventor is not active – the most conservative response – would radically reduce sample size, as most inventors patent only once (and would miss instances where inventors were constrained from patenting for some reason). I thus zero all cells when no inventor activity is recorded, and test 'blanking' in robustness checks. Second, we cannot be sure where inventors are when they are not actively patenting. Following the methodology set out in Agrawal et al (2006), I identify around 14% of the sample as likely movers.[12] In the main analysis I assume inventors do not move; I relax this in robustness checks.

Finally, we need to deal with individual-level heterogeneity, which if unobserved will bias the results. This is crucial as I need to distinguish effects of individuals' ethnicity from other personal characteristics. Ethnicity is time-invariant, so that individual fixed effects are not an

[11] TTWAs are designed to cover largely self-contained labour markets: 75% of those living in a given TTWA also work in the TTWA, and vice versa. TTWAs are thus a good approximation for local spatial economies and for city regions (Robson et al, 2006). Matching is done by postcode sector, which minimises observations lost through incomplete or mistyped postcode information (matching on full postcodes drops around 12% of observations. Matching on postcode sector drops 5.77%). I exclude inventors resident in Northern Ireland.

[12] 'Likely movers' are inventors with the same forename and surname, who patent in the same technology fields, in different TTWAs, at different points in time. This minimises the risk of false positives – identifying inventors who are movers who are not – but does not remove false negatives (identifying movers as non-movers).

option. I therefore deploy a widely-used alternative developed by Blundell et al (1995), using inventors' historic patenting behaviour. Blundell and colleagues argue that agents' capacity to innovate is largely explained by their knowledge at the point in which they enter a sample. With long enough time series data, pre-sample activity approximates an individual fixed effect. I replicate this estimator, using the pre-sample mean of inventors' patent counts as a human capital 'levels effect'.

I set the historic period as the 16 years between 1977 and 1992: around 23% of inventors in the sample period invent in this period, i.e. 40% of cells. I exclude inventors with no pre-sampling history – they may have been inactive or not in the labour force – and run the model on a reduced sample of 89,309 observations.[13] The new sample removes younger inventors and recent migrants: it may thus understate the true size of the UK's ethnic inventor population. This suggests the actual effects of diversity may be larger than those that I go on to identify.

5.2 Model

I estimate links between ethnicity and inventor behaviour using a modified knowledge production function, regressing counts of patenting activity on individual, group and area characteristics. For inventor *i* in area *j* and yeargroup *t*, I estimate:

[13] Fundamentally, I argue the reduced sample preferable to running a bigger sample of inventors for whom historic patenting information is ambiguous. Firm-level studies, in contrast, typically have information on exactly when agents enter/exit the market. I experiment with the full sample to check robustness, finding key variables and overall model fit are poor.

$$\text{PCOUNT}_{ijt} = a\text{INV}_i + b\text{DIV}_{jt} + \mathbf{CONTROLS_{jt}}\boldsymbol{c} + \text{P}_i + \text{U}_j + \text{YG}_t + \text{e}_i \quad (1)$$

Where PCOUNT is a simple count of the number of times an inventor engages in patenting during a given four-year period (which I refer to as 'inventor productivity'). My first variable of interest is INV, a dummy variable taking the value one if the inventor is a likely ethnic inventor. (I later extend the model replacing INV with a set of dummies for various co-ethnic groups.) My second key variable is DIV, the diversity of active inventors in a given TTWA and time period. For identity group *a* in area *j* in year *t*, the Index is given by:

$$\text{FRAC}_{jt} = 1 - \textstyle\sum_a [\text{SHARE}_{ajt}]^2 \quad (2)$$

Where SHARE is *a*'s share of the relevant population (here, all active inventors in a given area). The Index measures the probability that two individuals in an area come from different geographical origin or ethnic groups. Similar measures are used widely in the development literature, as well as some area-level studies (Ottaviano and Peri, 2006, Ottaviano and Peri, 2005, Alesina and La Ferrara, 2004, Easterley and Levine, 1997).

CONTROLS represents a vector of largely TTWA-level controls covering key spatial, economic, and demographic characteristics affecting relationships between INV and innovation, DIV and innovation or both. Unless otherwise stated, all controls are for the same 1993 – 2004 period as the patent data. I use aggregated LFS client file microdata to construct control variables.[14]

[14] I aggregate individual-level data to local authority-level averages, and then aggregate these to TTWA-level using postcode shares. Local Authority District (LAD) boundaries are not congruent with TTWA boundaries, so straightforward aggregation is not possible. Using the November 2008 National Postcode Sector Database (NSPD), I calculate the number of postcodes in each 2001 TTWA and in each of its constituent LADs. For each

Innovative activity and patenting are both spatially concentrated, reflecting benefits from agglomeration that may persist over time (Simmie et al., 2008). Co-ethnicity or diversity effects on patenting might then simply reflect agglomeration and path-dependence. I fit a dummy for primary urban areas, U, and fit log of population density to explore wider agglomeration effects. I also fit measures of 1981-84 area weighted patent stocks to control for historic asset effects, and experiment using different lags of the historic patent stocks control. Inventor demographic characteristics may be entirely explained by area demographic characteristics: for example, places with more diverse populations may produce more diverse inventor groups. I control for this by using area-level fractionalisation indices (and cross-check using migrant population shares). Human capital stocks are closely correlated with innovative activity (Romer, 1990) and may account for apparent ethnicity effects on patenting. Alongside the individual human capital 'levels effect', above, I fit areas' share of degree-holders with Science, Technology, Engineering and Mathematics (STEM) qualifications in the local working-age population.

I fit various further controls for precision. Patenting is known to be higher in 'knowledge-intensive' high-tech and manufacturing sectors, so I include measures of the share of workers employed in 'knowledge-intensive' manufacturing, following The Work Foundation's

---

TTWA, I then calculate constituent LADs' 'postcode shares'. Shares sum to one, and are used as weights to construct TTWA-level averages. *Example:* suppose a TTWA consists of parts of three LADs. The TTWA has 100 postcodes, 60 of which are in LAD_a, 30 in LAD_b and 10 in LAD_c. The relevant LAD weights are 0.6, 0.3 and 0.1 respectively. The TTWA-level average of variable x is given by (x)TTWA = 0.6*(x)_a + 0.3*(x)_b + 0.1*(x)_c.

definition of 'knowledge-intensive' firms (Brinkley, 2008).[15] Patenting activity is also vulnerable to sector-specific shocks, and the spike in software patenting since the mid-1990s is well-covered in the literature (Li and Pai, 2010). To account for this I fit a dummy for the IPC technology field 'electrical engineering and electronics'.[16] Patenting is likely to be lower in areas with a lot of entry-level jobs or areas of joblessness, so I include the share of workers in entry-level occupations and the share of long term unemployed as further controls. Summary statistics are given in Table 5.

5.3 Estimation

My panel exhibits excess zeroes (78%) and over-dispersion (the variance of PCOUNT is over 2.5 times the mean): the basic assumptions of the Poisson model are not met, leading to likely inefficient estimates (Greene, 1994). Diagnostic tests suggest a negative binomial is the most suitable replacement.[17] Against this, Angrist and Pischke (2009) argue that once raw coefficients are converted into marginal effects, non-linear modelling offers little over standard linear regression. I therefore fit the model with both negative binomial and OLS estimators.

[15] This adjusts OECD definitions for the UK context. The final list of 3-digit SIC sectors includes medium and high-tech manufacturing (pharmaceuticals, aerospace, computers and office machinery, electronic communications, software, other chemicals, non-electrical machinery, motors and transport equipment).

[16] I also experiment with a more precise information technology dummy , with similar results.

[17] Log-likelihood tests and AIC scores. I also experiment with zero-inflated models (ZIP and ZINB). Both perform well on diagnostic tests, although interpretation is extremely complex. Results available on request.

6. Regression analysis: results

6.1 All inventors

I first regress inventor productivity on the ethnic inventor dummy and inventor group diversity. Results are given in Table 6. Column 1 shows a bivariate regression for the main variables of interest only, column 2 adds controls and column 3 adds the human capital effect. Negative binomial results are presented as marginal effects at the mean, and show a significant log alpha term, confirming over-dispersion. Controls are generally of the expected size and sign.

Insert Table 6 about here

Fitted alone, ethnic status and inventor group composition have no significant effect on individual inventor productivity (column 1). When controls are added (column 2), both INV and DIV become positive. Coefficients get bigger, and in the OLS results DIV is now significant at 5%. Once individual-level heterogeneity is controlled for (column 3), overall model fit improves and the results change substantially. INV remains insignificant but its coefficient more than doubles, for both sets of models. For negative binomial models, the marginal effect of DIV is now 0.087, significant at 5%.

This implies that a 10-point increase in the inventor Fractionalisation Index – increasing active inventor diversity in Bristol to that in Oxford, for example – is linked to an average marginal effect of 10*(0.087) = 0.87 extra patents per inventor. For OLS models, diversity effects are slightly larger. DIV is 0.099, significant at 10%: a 10-point rise in inventor group diversity is associated with a 0.99 unit increase in expected patenting, or an extra patent per inventor. Interestingly, coefficients of *area population* diversity are negative in all models. To put this into perspective, effects of diversity on patent counts are smaller and/or weaker than human capital, whether the latter is measured at the area level or at individual level. This fits with the existing empirical evidence that diversity effects on innovation are generally fairly small, where they exist (see Section 2). For negative binomial models, for example, the marginal effect of STEM degrees is 0.304, significant at 5%. This suggests that a 10-point increase in the area's share of science graduates is linked to 3 extra patents per inventor. This is as expected given that patenting is concentrated in science and technology sectors. The marginal effect of the individual fixed effect 0.101, significant at 1%: past patenting activity is strongly linked to current patenting rates.

Insert Table 7 about here

Results for ONS ethnic groups function as a basic cross-check, and broadly confirm the main findings (Table 7). For negative binomial models, INV remains close to zero throughout; with controls and fixed effects the marginal effect of ethnic DIV is 0.125, significant at 5%. OLS results are not shown here: coefficient sizes and magnitudes are similar but none of the results is significant.

6.2 Urban areas

While the evidence review suggests that urban areas may 'amplify' ethnicity-innovation effects, I find little hard evidence of this. Coefficients for agglomeration are positive but not significant; the urban TTWA dummy is insignificant and close to zero. In order to identify the separate effects of urban location and urban density, I fit the two separately and then interact them.[18] Fitted separately, each is negative on inventor productivity (although marginally significant at best). Fitted together, each is positive – with a negative interaction effect, suggesting some diseconomies of agglomeration on inventor productivity in the largest conurbations. This is perhaps surprising given the emphasis on geographical proximity in the innovation literature. The UK context helps explain the discrepancy. Raw patent counts are highest in relatively small cities, notably Oxford and Cambridge. Conurbations, particularly London, are dominated by service sector activities where patenting is less likely to occur.

6.3 Co-ethnicity effects

To explore co-ethnic / diasporic group effects, I replace INV with dummies for each geographical origin zone. I run the model for all minority co-ethnic groups, taking UK-origin as the reference category. Results are interpreted as the marginal effect of being in one of these co-ethnic groups, relative to membership of the majority group of UK-origin inventors.

Insert Table 8 about here

Results are shown in Table 8. For simplicity I restrict the analysis to the five zones with the biggest numbers of ethnic inventors, as given in Table 2 (South Asia, Central Europe, East Asia, Southern Europe and Eastern Europe). I find significant positive effects of South Asian- and Southern European–origin inventors on expected patenting rates, and negative significant effects of East Asian-origin inventors, relative to UK-origin inventors. Specifically, marginal effects are 0.025 for South Asian inventors, significant at 10%, -0.037 (1%) for East Asian inventors; and 0.053 (10%) for Southern European inventors.

[18] Results are available on request.

The results have some echoes of US studies, but point up important differences.[19] For example, the South Asian result is likely driven by the strong historic connections between the UK and India, Pakistan, and Bangladesh, and the presence of large South-Asian origin communities in Britain. The Southern European result is likely to reflect the UK's large shares of inventors with Spanish, Italian or Portuguese backgrounds (Table 1). The East Asian result is in stark contrast to US research. It may reflect the lack of strong diasporas in the UK outside Hong Kong-origin Chinese, and the different circumstances behind recent community formation in the US (economic migration of skilled workers) and the UK (handover of Hong Kong to China between 1984 and 1997).

7. Robustness checks

I conduct a number of robustness tests on both the pooled and co-ethnic group models. Results for the former are given in Table 9; results for the latter are available on request, but are very similar.

Insert Table 9 about here

I first run some basic specification checks. In Table 9, Column 2 includes a lagged dependent variable to control for effects of past patenting within the sample. Diversity effects persist: coefficients are now rather smaller but also more precise, with DIV significant at 1%. Column 3 fits the model without London – a city with high levels of cultural diversity and

[19] I conduct cross-checks using more tightly-defined country of origin groups, confirming these results. Findings are available on request.

relatively low levels of patenting per head of population (Wilson, 2007). Again, diversity effects persist in the negative binomial specification (significant at 5%). Column 4 fits the TTWA share of PhD-holders as an alternative area-level human capital control. This is strongly associated with inventor productivity, and dominates DIV in both model specifications. Places attractive to PHDs may also attract a diverse inventor group, via some other factor – such as a 'tolerant' milieu as suggested by Florida (2002). Alternatively, high-patenting PHDs are themselves ethnic inventors, as suggested by US studies on star scientists (Stephan and Levin, 2001, Chellaraj et al., 2005). In this case, diversity is the fundamental driver and the PhD variable is a so-called 'bad control' (Angrist and Pischke, 2009). Further research is needed here, perhaps with a subset of academic inventors. In the meantime, some caution is needed in interpreting the diversity result as causal.

If my existing specification is not fully capturing inventors' human capital, resulting omitted variable bias will overstate effects of DIV. I test this using an alternative human capital effect capturing intellectual range: I identify 'generalists' as inventors patenting across at least two technology fields. In Table 9, column 5 fits a generalist dummy alone; marginal effects of DIV fall from 0.087 to 0.05, 10% significance. Column 6 fits both human capital controls. This improves the size and strength of the DIV effect, and improves model fit. In both cases, diversity effects survive.

Inventor diversity effects might also collapse to simple size effects, not least because Fractionalisation Indices tend to be highly correlated with group population shares (in this case, the pairwise correlation between the two is 0.8039). To test this, I replace the Fractionalisation Index of inventors with the share of ethnic inventors in the local inventor population. Column 7 in Table 9 fits the size control alone; column 8 fits size and DIV together; column 9 interacts the two. Ethnic inventor shares are insignificant in all cases, while the interaction term is negative insignificant. This suggests that the overall diversity of inventors, rather than an aggregation of ethnic inventors, drives the main results.

Finally, innovative activity is also spatially concentrated, and these concentrations tend to persist over time. If the historic patent stocks term in the main model is mis-specified, agglomeration and path-dependence will not be adequately controlled for. To test for this I plug a range of pre-sample historic patent counts into the main model. Column 10 shows results for the most conservative specification (when the lag is dropped to the four-year period before the sample). Coefficients of DIV fall to 0.067, significant at 10%. This suggests that historic area-level characteristics help explain some of the inventor diversity effect – but do not eliminate it.

7.2 Further checks

I conduct two further tests for more structural issues. First, I reconstruct my sample by blanking all inventor-yeargroup cells when an inventor is not patenting. As discussed in Section X, this is a more conservative way of treating inventors when they are not active, and will capture any measurement error introduced by my choice of zeroing. My identification strategy depends on using inventors' historic patenting activity, so blanking out non-activity has the effect of restricting the sample to inventors who patent more than once. I thus compare estimates for the set of multiple inventors across two different samples, one with zeroed and one with missing observations for non-activity. Results show that estimates for the two sub-samples are identical.[20] Sample construction therefore has no effect on my main findings.

Second, I relax my assumption of immobile inventors. Around 14% of inventors as likely movers across TTWAs: allowing inventors to move raises the possibility of positive selection into high-patenting regions, creating upwards bias on DIV. I test for endogeneity using lags of DIV as instruments. Diagnostics suggest a two-yeargroup (eight year) lag is preferred. Following Cameron and Trivedi (2009) I estimate using a two-step structural model to give consistent estimates at the second stage. Results are given in Table 10. The instrument passes first stage tests assessed using 2SLS (column 1). Second-stage results show much larger coefficients of DIV, although less precise (column 2). Overall, the IV findings confirm that my main results are not sensitive to inventor mobility.

8. Distributional impacts

The analysis so far has ignored distributional effects – that is, specific impacts of ethnic inventors on majority inventors. I explore two distributional issues here.

I first look at native outflows, in which UK-origin inventors physically leave an area after minorities arrive (Borjas, 1994). Using the subset of likely movers, I conduct exploratory logit regressions to identify factors influencing mover status. Results suggest individual human capital has a substantial, significant positive link to mover status. By contrast,

[20] Results available on request.

coefficients for areas' share of migrant inventors are much smaller and statistically insignificant.

Next, I explore 'resource crowd-out' – a potentially more serious issue. For instance, a given majority inventor may benefit from ethnic inventors via production complementarities, or may 'lose' from disbenefits such as lower trust or communications difficulties. Even if there are human capital externalities at the group level, ethnic inventors might crowd out majority inventors from relevant jobs and resources, such as space in R&D labs; or diaspora benefits might only be accessible to group members (Borjas and Doran, 2012). This will affect the composition of overall patenting at area level. At the extreme, increases in area-level patenting might be wholly explained by 'ethnic' patents. Conversely, there might be multiplier effects from ethnic inventors to majority group inventors.

I test for both forms of resource crowd-out (Table 11). At the individual level, I first re-run model (1) for majority inventors only. Results are given in the first panel. The marginal effect of DIV on majority inventor productivity is positive significant, but smaller and weaker than on all inventors. I then re-run (1) but fit INV as a majority inventor dummy. Results are given in the second panel. Majority status alone has no effect on inventor productivity (columns 1 and 2), but when interacted with inventor diversity, I find a negative significant result for majority inventors in diverse areas (column 3). Unlike the previous test, this suggests that while inventor diversity brings benefits, majority inventors in diverse inventor communities lose out.

To explore area-level effects, I draw on recent work by Card (2005), Kerr and Lincoln (2010) and Faggio and Overman (2011). I assemble a panel of TTWA-level weighted patent counts for 1993-2004. I define ‘ethnic’ patents as patents with at least one ethnic inventor; all other patents are ‘majority’ patents. Following Faggio and Overman (2011), I then regress the percentage change in total weighted patents during the period on the percentage change in ethnic patents. For TTWA *j* I estimate:

$$\Delta TPATENTS_j = a + b\Delta EPATENTS_j + \mathbf{CONTROLS}\boldsymbol{c}_{jtbase} + e_j \quad (3)$$

Where:

$$\Delta TPATENTS_j = TPATENTS_{j2004} - TPATENTS_{j1993} / TPATENTS_{j1993} \quad (4)$$

And $\Delta EPATENTS_j$ is assembled similarly. **CONTROLS** is a vector of area-level controls for the base period 1993. The coefficient of interest is *b*. As explained by Card (2005), if estimates of *b* are less than one, increases in ethnic patenting lead to a smaller increase in overall patenting, implying some crowd-out of majority patenting by ethnic inventors. Estimates of *b* larger than one imply multiplier effects; if *b* is equal to one, there are no distributional impacts either way.

OLS results are given in the third panel of Table 11. The simplest specifications of (4) suggest some crowd-out, with *b* estimated at 0.199 and 0.259, significant at 1%. However, *b* becomes insignificant once controls and standard errors clustered on TTWAs are introduced (column 4). An alternative specification using shifts in TTWAs’ technology field shares delivers very similar results (column 5). This suggests there is little evidence of crowd-out –

although as these are correlations rather than causal effects, results should be interpreted carefully.[21]

9. Conclusions

In recent years there has been growing academic and policy interest in links between immigration, ethnic diversity and innovation. This paper looks at the role of ethnic inventors on innovative activity in the UK, using a new 12-year panel of patents microdata and a novel name classification system. I uncover some distinctive features of the UK ethnic inventor community, and have been able to explore a number of potential 'ethnicity-innovation' channels – individual positive selection, externalities from diasporic groups and from the cultural diversity of inventor communities, as well as 'amplifying' effects of urban environments. The research is one of very few studies to explore these links, and as far as I am aware is the first of its kind outside the US.

The descriptive analysis suggests that UK's ethnic inventor community has some important commonalities with the US – with large South and East-Asian origin groups – and some key differences, reflecting proximity to Continental Europe and recent immigration trends. Ethnic inventors are spatially clustered, but not all high-patenting regions have diverse inventor communities.

Regressions suggest that 'ethnic inventor' status has no significant effect on inventor patenting rates once other factors are controlled for. Conversely some diasporic groups, and

[21] I experiment with lags of ethnic patents as an instrument, but none pass the required first-stage tests.

group cultural composition, have small positive effects on inventor productivity. Effects on 'majority' inventors are unclear: there are some indications of individual-level crowd-out, but not at area level. Although patenting activity is very spatially clustered in the UK, in contrast to the wider literature, I find little evidence that urban environments improve individuals' patenting activity once other individual and area-level controls are taken into account. Overall, the results suggest that ethnic inventors are a net positive for patenting in the UK. My results are also likely to understate the true effects of diversity and co-ethnic groups.

The results also point to important cross-country differences in ethnic inventor behaviour and outcomes. Notably, US ethnic inventor communities have been shaped by Cold War science research, which have attracted very large numbers of skilled workers into a small number of locations (Saxenian, 2006). By contrast, recent 'calls' for migrant workers in the UK since the mid-20th century have been largely focused on less skilled occupations, although policy is now becoming more skill-biased. Results may also reflect culturally distinctive US attitudes to entrepreneurship, as evidenced by sociological studies of Jewish and Afro-Caribbean migrant communities in New York and London (Gordon et al., 2007), and by the complex interplay between class, skills, resources and attitudes that influence real-world entrepreneurial behaviour (Basu, 2002).

There are three important caveats to these results. First, diversity and diaspora effects are relatively small – human capital and patent field / industry effects are more important determinants of inventors' productivity. This is intuitive, and echoes much of the existing literature (see above). Second, working with inventor data presents a number of potential measurement error challenges. Most seriously, my data only allows a fuzzy identification of ethnic inventors and diasporic groups. Third, although the results survive a number of robustness checks, alternative measures of area-level human capital weaken effects of DIV, suggesting caution in interpreting causality as one-way.

The results have implications for the current UK government's migration policies. Net immigration is one of the main factors behind the growth of ethnic inventor communities in the UK: a phenomenon which appears to raise rates of innovation through a combination of diversity and diaspora effects, with no hard evidence of negative distributional effects on native inventors. The current migration cap places restrictions on skilled immigration from outside Europe – so will likely put constraints on innovative activity, leading to welfare losses both to the UK and to UK-born workers. Similar welfare losses may arise from proposed restrictions on post-study routes to work for non-EU students.

The paper leaves a number of questions for future research. This could explore social networks, co-ethnicity and geographical location in more detail – via analysis of patent citations and international co-invention / co-patenting. Within the UK, data offering better identification of ethnic and migrant inventors, in particular recent immigrants, would provide a clearer picture of current developments. Alternatively, qualitative methods could shine further light on migrant and diaspora dynamics. Further work is needed on academic inventors, and on the relative contribution of majority and ethnic PHDs to patenting. Finally, the analysis should be extended to other EU countries.

LIST OF TABLES

Table 1. UK-resident inventors: 30 biggest CEL subgroups, 1993-2004.

CEL subgroup	Freq.	%	Cumulative %
ENGLISH	86,118	69.17	69.17
CELTIC	10,653	8.56	77.73
SCOTTISH	6,557	5.27	82.99
IRISH	3,583	2.88	85.87
WELSH	2,523	2.03	87.9
INDIAN HINDI	1,255	1.01	88.91
GERMAN	1,205	0.97	89.87
ITALIAN	975	0.78	90.66
FRENCH	958	0.77	91.43
CHINESE	920	0.74	92.16
POLISH	886	0.71	92.88
OTHER MUSLIM	793	0.64	93.51
OTHER EUROPEAN	665	0.53	94.05
HONG KONGESE	588	0.47	94.52
GREEK	574	0.46	94.98
PAKISTANI	551	0.44	95.42
SIKH	500	0.4	95.82
SPANISH	438	0.35	96.18
VIETNAMESE	427	0.34	96.52
JEWISH	351	0.28	96.8
PORTUGUESE	326	0.26	97.06
JAPANESE	293	0.24	97.3
EAST ASIAN & PACIFIC	263	0.21	97.51
DANISH	216	0.17	97.68
OTHER SOUTH ASIAN	209	0.17	97.85
SRI LANKAN	209	0.17	98.02
DUTCH	207	0.17	98.19
TURKISH	198	0.16	98.34
SWEDISH	191	0.15	98.5
RUSSIAN	138	0.11	98.61

Source: ONOMAP/KITES-PATSTAT.

Notes:

1) ‘OTHER MUSLIM’ subgroup includes CEL name types ‘BALKAN MUSLIM’, ‘MALAYSIAN MUSLIM’, ‘MUSLIM INDIAN’, ‘SUDANESE’, ‘WEST AFRICAN MUSLIM’, ‘OTHER MUSLIM’ (SMALLER MIDDLE EASTERN COUNTRIES, N/AFRICAN COUNTRIES, CENTRAL ASIAN REPS)
2) 'JEWISH' includes CEL name types ‘JEWISH / ASHKENAZI’, ‘SEPHARDIC JEWISH’
3) ‘EAST ASIAN AND PACIFIC’ includes CEL name types ‘BURMESE’, ‘CAMBODIAN’, ‘FIJIAN’, ‘HAWAIIAN’, ‘LAOTIAN’,‘MAORI’, ‘MAURITIAN’, ‘POLYNESIAN’, ‘SAMOAN’, ‘SINGAPOREAN’, ‘SOLOMON ISLANDER’, ‘SOUTH EAST ASIAN’ , ‘THAI’, ‘TIBETIAN’, ‘TONGAN’, ‘TUVALUAN’, ‘EAST ASIAN & PACIFIC OTHER’
4) ‘OTHER SOUTH ASIAN’ includes CEL name types ‘ASIAN CARIBBEAN’, ‘BENGALI’, ‘BHUTANESE’, ‘GUYANESE ASIAN’, ‘KENYAN ASIAN’, ‘NEPALESE’, ‘PARSI’, ‘SEYCHELLOIS’, ‘SOUTH ASIAN’, ‘TAMIL’

Table 2. UK-resident inventors: geographical origin and ONS ethnic groups, 1993-2004.

Probable geographical area of origin	**Freq.**	**%**	**Cumulative %**
BRITISH ISLES	109,429	87.89	87.89
SOUTH ASIA	3,074	2.47	90.36
CENTRAL EUROPE	3,035	2.44	92.8
EAST ASIA	2,557	2.05	94.85
SOUTHERN EUROPE	2,394	1.92	96.78
EASTERN EUROPE	1,395	1.12	97.9
MIDDLE EAST	1,060	0.85	98.75
NORTHERN EUROPE	606	0.49	99.24
REST OF WORLD	568	0.46	99.70
AFRICA	324	0.26	99.96
CENTRAL ASIA	31	0.02	99.98
AMERICAS	29	0.02	100
Probable ethnic group in 1991 Census categories			
WHITE	.	94.28	94.28
ANY OTHER ETHNIC GROUP	.	1.76	96.04
INDIAN	.	1.69	97.73
CHINESE	.	1.41	99.14
PAKISTANI	.	0.54	99.68
BLACK - AFRICAN	.	0.24	99.92
BANGLADESHI	.	0.08	100
BLACK - CARIBBEAN	.	0	100
BLACK - OTHER	.	0	100

Source: ONOMAP/KITES-PATSTAT.
Notes: Ethnic groups typology taken from 1991 Census to allow comparability pre and post-2001. Frequencies have been supressed to avoid disclosure.

Table 3. Ethnic inventor Location Quotients, 1993-2004. Top 40 areas.

LQ	TTWA name	TTWA type
2.372	Crawley	Primary Urban
1.989	Southampton	Primary Urban
1.703	London	Primary Urban
1.414	Oxford	Primary Urban
1.394	Cambridge	Primary Urban
1.375	Dundee	Primary Urban
1.304	Oban	N Scotland rural
1.266	Guildford & Aldershot	Primary Urban
1.252	Swindon	Primary Urban
1.216	St Andrews & Cupar	N Scotland rural
1.213	Edinburgh	Primary Urban
1.180	Pembroke & Tenby	Welsh rural
1.180	Colchester	Primary Urban
1.162	Carlisle	N England rural
1.139	Bude & Holsworthy	SW England rural
1.122	Aberdeen	Primary Urban
1.101	Holyhead	Welsh rural
1.062	Brighton	Primary Urban
1.044	Lancaster & Morecambe	N England rural
1.024	Bedford	Primary Urban
1.005	Livingston & Bathgate	N Scotland rural
1.000	Cardiff	Primary Urban
0.995	Glasgow	Primary Urban
0.988	Inverness & Dingwall	N Scotland rural
0.981	Lanarkshire	Primary Urban
0.980	Newcastle & Durham	Primary Urban
0.955	Birmingham	Primary Urban
0.953	Haverfordwest & Fishguard	Welsh rural
0.941	York	Primary Urban
0.940	Leicester	Primary Urban
0.938	Reading & Bracknell	Primary Urban
0.932	Wycombe & Slough	Primary Urban
0.917	Wirral & Ellesmere Port	Primary Urban
0.898	Leeds	Primary Urban
0.897	Newbury	SW England rural
0.893	Louth & Horncastle	Rest England rural
0.886	Liverpool	Primary Urban
0.876	Canterbury	Rest England rural
0.875	Margate, Ramsgate & Sandwich	Rest England rural
0.872	Harlow & Bishop's Stortford	Rest England rural

Source: ONOMAP/KITES-PATSTAT/ONS.

Note: TTWAs use 2001 boundaries. ‘Primary urban’ TTWAs contain an urban core with at least 125,000 people. TTWAs with fewer than 10 inventors suppressed.

Table 4. TTWAs' weighted patent stocks, 1993-2004. Top 40 areas.

Weighted patent count	TTWA name	TTWA type
1697.14	London	Primary Urban
1155.59	Cambridge	Primary Urban
719.36	Oxford	Primary Urban
705.62	Harlow & Bishop's Stortford	Rest England rural
531.69	Manchester	Primary Urban
489.87	Guildford & Aldershot	Primary Urban
483.41	Southampton	Primary Urban
440.96	Bristol	Primary Urban
428.15	Reading & Bracknell	Primary Urban
416.01	Crawley	Primary Urban
379.21	Ipswich	Primary Urban
365.63	Swindon	Primary Urban
342.90	Wycombe & Slough	Primary Urban
341.67	Stevenage	Primary Urban
312.93	Newcastle & Durham	Primary Urban
309.40	Wirral & Ellesmere Port	Primary Urban
301.75	Leicester	Primary Urban
289.82	Birmingham	Primary Urban
260.66	Nottingham	Primary Urban
223.87	Leeds	Primary Urban
218.49	Edinburgh	Primary Urban
213.60	Worcester & Malvern	Primary Urban
183.83	Margate, Ramsgate & Sandwich	Rest England rural
181.10	Coventry	Primary Urban
169.36	Bedford	Primary Urban
167.98	Luton & Watford	Primary Urban
165.09	Cardiff	Primary Urban
163.87	Glasgow	Primary Urban
161.37	Warwick & Stratford-upon-Avon	Rest England rural
161.20	Warrington & Wigan	Primary Urban
152.70	Hull	Primary Urban
148.04	Derby	Primary Urban
147.14	Aberdeen	Primary Urban
138.16	Portsmouth	Primary Urban
136.70	Milton Keynes & Aylesbury	Primary Urban
130.99	Middlesbrough & Stockton	Primary Urban
121.67	Chelmsford & Braintree	Primary Urban
121.35	Chester & Flint	Welsh rural
118.13	Northampton & Wellingborough	Primary Urban
113.95	Maidstone & North Kent	Primary Urban

Source: KITES-PATSTAT/ONS.

Note: TTWAs use 2001 boundaries. 'Primary urban' TTWAs contain an urban core with at least 125,000 people. Patents are weighted by number of inventors, *not* area population.

Table 5. Summary statistics.

Variable	N	Mean	SD	Min	Max
Inventor patent count / 4-year period	89312	0.114	0.694	0	25
Inventors' ave patent count, pre-1993	89312	0.405	0.351	0.286	11.143
Inventor likely techfield mover	89312	0.256	0.437	0	1
Inventor likely TTWA mover	89312	0.143	0.35	0	1
Inventor is UK geog. origin	89312	0.937	0.243	0	1
Inventor is foreign geog. origin	89312	0.063	0.243	0	1
Inventor African origin	89312	0.002	0.041	0	1
Inventor Americas origin	89312	0.000	0.013	0	1
Inventor Central Asia origin	89312	0.000	0.018	0	1
Inventor Central Europe origin	89312	0.012	0.107	0	1
Inventor rest of world origin	89312	0.003	0.058	0	1
Inventor East Asian origin	89312	0.007	0.084	0	1
Inventor East Europe origin	89312	0.007	0.086	0	1
Inventor Middle East origin	89312	0.006	0.075	0	1
Inventor Northern Europe origin	89312	0.003	0.052	0	1
Inventor South Asian origin	89312	0.015	0.123	0	1
Inventor South European origin	89312	0.007	0.086	0	1
Frac. Index, geog. origin groups	89312	0.209	0.118	0	0.612
Inventor is white ethnicity	89312	0.97	0.172	0	1
Inventor is minority ethnic	89312	0.03	0.172	0	1
Inventor Black Caribbean	89312	0	0.01	0	1
Inventor Black African	89312	0.002	0.04	0	1
Inventor Indian	89312	0.012	0.107	0	1
Inventor Pakistani	89312	0.003	0.052	0	1
Inventor Chinese	89312	0.004	0.064	0	1
Inventor other ethnic group	89312	0.01	0.099	0	1
Frac. Index, ethnic groups	89312	0.108	0.066	0	0.449
TTWA Frac Index, geog. groups	89309	0.225	0.142	0	0.528
TTWA Frac Index, ethnic groups	89309	0.169	0.141	0	0.459
% graduates	89309	0.238	0.051	0.106	0.362
% graduates with STEM degrees	89309	0.121	0.032	0.041	0.196
% graduates with PhDs	89309	0.007	0.005	0	0.029
% employed hi-tech manufacturing	89309	0.027	0.014	0	0.194
% employed medium-tech m’facturing	89309	0.046	0.023	0	0.135
% in entry level occupations	89309	0.338	0.049	0.25	0.667
% unemployed >=12 months	89309	0.016	0.012	0	0.08
log(population density)	89309	6.605	1.053	2.06	8.359
Electronics patent	89312	0.009	0.093	0	1
TTWA weighted patent count	89312	493.094	578.301	0	1888.03
TTWA weighted patents, 1981-84	88726	144.814	201.789	0.25	613.859

Source: KITES-PATSTAT/ONS/LFS
Note: Area-level controls not available for all TTWAs.

Table 6. Patent counts, geographical origin zones, 1993-2004.

	Negative Binomial			OLS		
	(1)	(2)	(3)	(1)	(2)	(3)
Ethnic inventor, geog.	-0.000 (0.011)	0.004 (0.008)	0.009 (0.007)	-0.002 (0.011)	0.004 (0.011)	0.011 (0.010)
Frac Index of inventors, geog. origin groups	-0.061 (0.101)	0.079 (0.050)	0.087** (0.042)	-0.055 (0.088)	0.119** (0.058)	0.099* (0.055)
Frac Index, TTWA country of birth		-0.203* (0.110)	-0.140* (0.085)		-0.137 (0.127)	-0.079 (0.115)
% STEM degrees, TTWA		0.372** (0.176)	0.304** (0.147)		0.302 (0.292)	0.334 (0.278)
Log of TTWA population density		0.005 (0.008)	0.005 (0.007)		0.006 (0.010)	0.008 (0.009)
Area weighted patents, 1981-84		-0.000 (0.000)	-0.000 (0.000)		-0.000 (0.000)	-0.000 (0.000)
% hi-tech mf empl, OECD defn.		-0.159 (0.281)	-0.111 (0.226)		-0.166 (0.385)	-0.245 (0.367)
% medium-tech mf, OECD defn.		0.048 (0.172)	0.051 (0.134)		0.120 (0.240)	0.093 (0.216)
% entry-level occupations		0.042 (0.123)	0.113 (0.106)		0.084 (0.166)	0.149 (0.154)
% unemployed >=12 months		-0.313 (0.441)	-0.000 (0.354)		-1.211 (0.747)	-0.934 (0.719)
Electronics / OST7 type 1 patents		2.074*** (0.132)	1.697*** (0.176)		2.356*** (0.139)	2.305*** (0.135)
Urban TTWA		-0.018* (0.015)	-0.021* (0.015)		-0.024 (0.019)	-0.028 (0.017)
Human capital effect			0.101*** (0.007)			0.266*** (0.036)
ln(alpha) / Constant	2.991*** (0.052)	2.683*** (0.063)	2.491*** (0.069)	0.196*** (0.010)	0.122 (0.107)	-0.034 (0.105)
Observations	89312	88726	88726	89312	88726	88726
Log-likelihood / F	-25328.463	-24379.554	-23859.107	76.283	52.523	50.226
Chi^2 fit (Wald) / R^2	376.947	3520.345	2693.200	0.007	0.107	0.125

Source: KITES-PATSTAT/ONS/LFS

Notes: Notes: All models use time dummies. Heteroskedasticity and autocorrelation-robust standard errors clustered on TTWA. Negative binomial results: except for ln(alpha) term, coefficients are marginal effects at the mean. OLS results are raw coefficients. * = significant at 10%, ** 5%, *** 1%.

Table 7. Patent counts, all inventors, ONS ethnic groups.

Negative binomial

Individual patent counts	**(1)**	**(2)**	**(3)**
Ethnic inventor, ONS minority ethnic group	-0.006	-0.000	0.000
	(0.016)	(0.014)	(0.012)
Frac Index of inventors, ONS ethnic groups	-0.165	0.101	0.125**
	(0.145)	(0.067)	(0.056)
Controls	N	Y	Y
Human capital effects	N	N	Y
Observations	89312	88726	88726
Log-likelihood	-25319.277	-24386.644	-23864.136
Chi^2 goodness of fit statistic (Wald)	414.921	2706.003	2426.458

OLS

Individual patent counts	**(1)**	**(2)**	**(3)**
Ethnic inventor, ONS minority ethnic group	-0.010	-0.002	0.003
	(0.015)	(0.014)	(0.013)
Frac Index of inventors, ONS ethnic groups	-0.155	0.123	0.097
	(0.131)	(0.082)	(0.077)
Controls	N	Y	Y
Human capital effects	N	N	Y
Observations	89312	88726	88726
F-statistic	75.337	54.477	58.197
R^2	0.007	0.107	0.125

Source: KITES-PATSTAT/ONS/LFS.
Notes: All models use time dummies. Controls fitted: log of population density, % STEM degrees, % employed in knowledge-intensive manufacturing, fractionalisation index of area birth country groups, % entry-level occupations, % long term unemployed, urban TTWA dummy. Heteroskedasticity and autocorrelation-robust standard errors clustered on TTWA. Negative binomial models show marginal effects at the mean. * = significant at 10%, ** 5%, *** 1%.

Table 8. Inventor groups, negative binomial results.

Inventor patent count	Marginal effect
Africa origin	-0.037* (0.022)
Americas origin	0.176 (0.166)
Central Asia origin	0.045 (0.055)
Central Europe origin	-0.003 (0.014)
Diasporic origin	-0.019 (0.014)
East Asia origin	-0.037*** (0.007)
Eastern Europe origin	0.032 (0.034)
Middle East origin	-0.008 (0.025)
Northern Europe origin	0.001 (0.045)
South Asia origin	0.025* (0.015)
Southern Europe origin	0.053* (0.040)
Frac Index of inventors, geog. origin groups	0.087** (0.042)
Controls	Y
Observations	88726
Log-likelihood	-23843.642
Chi-squared	4438.933

Source: KITES-PATSTAT/ONS/LFS
Notes: all models use time dummies. Robust standard errors clustered on TTWA. Controls fitted: log of population density, % STEM degrees, % employed in knowledge-intensive manufacturing, fractionalisation index of ONS ethnic groups, % entry-level occupations, % long term unemployed, urban TTWA dummy. Coefficients are marginal effects at the mean. * = significant at 10%, ** 5%, *** 1%.

Table 9. Robustness checks. Negative binomial results.

Individual patent counts	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ethnic inventor, geographic origin	0.009 (0.007)	0.007 (0.007)	-0.000 (0.001)	-0.002 (0.001)	0.003 (0.004)	0.005 (0.004)	0.009 (0.007)	0.009 (0.007)	0.008 (0.007)	0.008 (0.007)
Frac Index of inventors, geog. origin groups	0.087** (0.042)	0.046 (0.039)	0.016*** (0.006)	0.016*** (0.006)	0.050* (0.028)	0.055** (0.027)		0.108*** (0.041)	0.191** (0.080)	0.067* (0.040)
% with PhDs in TTWA		2.649*** (0.504)								
#times inventor patents in previous YG within sample			0.053*** (0.002)	0.057*** (0.002)						
Alt human capital effect (patents in >1 IPC7 field)					0.217*** (0.010)	0.184*** (0.009)				
% ethnic inventors / all							0.068 (0.145)	-0.058 (0.138)	0.060 (0.121)	
Frac index * % ethnic inventors									-0.676* (0.345)	
Area historic weighted stock of patents, 1989-1992										0.000 (0.000)
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
HC effect	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
Include London?	Y	Y	Y	N	Y	Y	Y	Y	Y	Y
Observations	88726	88726	88726	75571	88726	88726	88726	88726	88726	89268
Log-likelihood	-23859.107	-23821.523	-16507.273	-21524.746	-22138.191	-21926.052	-23868.208	-23858.221	-23851.433	-24030.991
Chi^2 fit statistic (Wald)	2693.200	2181.073	4008.364	2095.403	3670.001	5323.670	3064.329	2830.487	3853.584	2865.519

Source: KITES-PATSTAT/ONS/LFS
Notes:-All models use time dummies. Controls fitted: log of population density, % STEM degrees, % employed in knowledge-intensive manufacturing, fractionalisation index of area birth country groups, % entry-level occupations, % long term unemployed, urban TTWA dummy. Heteroskedasticity and autocorrelation-robust standard errors clustered on TTWA. Coefficients are marginal effects at the mean. * = significant at 10%, ** 5%, *** 1%.

Table 10. Instrumental variable results, two-period lags of inventor fractionalisation index.

Inventor patent counts	(1) first stage	(2) second stage
Ethnic inventor, geog origin	0.008	0.089
	(0.010)	(0.077)
Frac Index of inventors, geog. origin groups	0.479**	4.789***
	(0.200)	(0.904)
Residual		-3.921***
		(0.926)
Constant	-0.049	-3.766***
	(0.121)	(0.512)
lnalpha		2.489***
Constant		(0.039)
Controls	Y	Y
Observations	88726	89312
F	44.852	.
R^2	0.124	.
Kleibergen-Paap under-identification test	19.236	.
Kleibergen-Paap weak identification test	30.817	.

Source: KITES-PATSTAT/ONS/LFS
Notes: all models use time dummies. First stage results are derived using 2SLS and are illustrative only. Second stage results are marginal effects from negative binomial estimation, derived using two-step structural equations. Standard errors are bootstrapped, 50 reps. Controls fitted for both models: log of population density, % STEM degrees, % employed in knowledge-intensive manufacturing, fractionalisation index of ONS ethnic groups, % entry-level occupations, % long term unemployed, urban TTWA dummy. Coefficients are marginal effects at the mean. * = significant at 10%, ** 5%, *** 1%.

Table 11. Distributional effects: individual and area-level analysis.

Individual level

Individual patent counts	**(1)**	**(2)**	**(3)**
UK inventor	-0.010 (0.008)	-0.009 (0.007)	0.027*** (0.008)
Frac Index of inventors, geog. origin groups		0.087** (0.042)	0.253*** (0.077)
UK * Frac Index			-0.172*** (0.056)
Controls	Y	Y	Y
Observations	88726	88726	88726
Log-likelihood	-23870.231	-23859.107	-23852.425
Chi^2 fit statistic (Wald)	3421.238	2693.200	2866.909

Source: KITES-PATSTAT/ONS/LFS
Notes: All models use time dummies. Controls fitted: log of population density, % STEM degrees, % employed in knowledge-intensive manufacturing, fractionalisation index of area birth country groups, % entry-level occupations, % long term unemployed, urban TTWA dummy. Heteroskedasticity and autocorrelation-robust standard errors clustered on TTWA. Coefficients are marginal effects at the mean. * = significant at 10%, ** 5%, *** 1%.

Area level

% change in total weighted patents, 1993-2004	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
% change in weighted ethnic patents, 1993-2004	0.199*** (0.065)	0.259*** (0.066)	0.248*** (0.068)	0.248 (0.177)	0.259 (0.178)
Controls	N	Y	Y	Y	Y
OST7 technology field dummies	N	N	Y	Y	N
HAC standard errors	N	N	N	Y	Y
Observations	220	220	210	210	206
F-statistic	9.299	1.467	3.646	1.144	0.966
R^2	0.041	0.041	0.141	0.141	0.151

Source: KITES-PATSTAT/ONS/LFS
Notes: All models use time dummies. Controls fitted: log of population density, % STEM degrees, % employed in knowledge-intensive manufacturing, % migrant working-age population, % entry-level occupations, % long term unemployed, urban dummy. Technology field dummies cover OST7 fields 1 -6: electrical engineering and electronics; instruments; chemicals and materials; pharmaceuticals and biotechnology; industrial processes; mechanical engineering, machines and transport. Consumer goods and civil engineering patents are used as the reference category. * = significant at 10%, ** 5%, *** 1%.

REFERENCESStuff

AGRAWAL, A., COCKBURN, I. & MCHALE, J. 2006. Gone but not forgotten: knowledge flows, labor mobility, and enduring social relationships. *Journal of Economic Geography,* 6**,** 571-591.

AGRAWAL, A., KAPUR, D. & MCHALE, J. 2008. How do spatial and social proximity influence knowledge flows? Evidence from patent data. *Journal of Urban Economics,* 64**,** 258-269.

AGRAWAL, A., KAPUR, D., MCHALE, J. & OETTL, A. 2011. Brain Drain or Brain Bank? The impact of skilled emigration on poor country innovation. *Journal of Urban Economics,* 69**,** 43-55.

ALESINA, A. & LA FERRARA, E. 2004. Ethnic Diversity and Economic Performance. *NBER Working Paper 10313.* Cambridge, MA: NBER.

ANGRIST, J. & PISCHKE, J.-S. 2009. *Mostly Harmless Econometrics,* Princeton, Princeton University Press.

ARCHIBUGI, D. & IAMMARINO, S. 2002. The globalization of technological innovation: definition and evidence. *Review of International Political Economy,* 9**,** 98–122.

ASPINALL, P. 2009. The Future of Ethnicity Classifications. *Journal of Ethnic and Migration Studies,* 35**,** 1417-1435.

AUDRETSCH, D. & FELDMAN, M. 1996. R&D Spillovers and the Geography of Innovation and Production. *American Economic Review,* 86**,** 630-640.

BASU, A. 2002. The Interaction between Culture and Entrepreneurship in London's Immigrant Businesses. *International Small Business Journal,* 20**,** 371-393

BASU, A. 2004. Entrepreneurial aspirations among family business owners: An analysis of ethnic business owners in the UK. *International Journal of Entrepreneurial Behaviour & Research,* 10**,** 12-33.

BERLIANT, M. & FUJITA, M. 2009. The Dynamics of Knowledge Diversity and Economic Growth. *56th Annual North American Meeting, Regional Science Association International.* San Francisco.

BLUNDELL, R., GRIFFITH, R. & VAN REENEN, J. 1995. Dynamic Count Data Models of Technological Innovation. *The Economic Journal,* 105**,** 333-344.

BONACICH, E. 1973. A Theory of Middleman Minorities. *American Sociological Review,* 38**,** 583-594.

BORJAS, G. 1987. Self-Selection and the Earnings of Immigrants. *American Economic Review,* 77**,** 531-53.

BORJAS, G. 1994. The Economics of Immigration. *Journal of Economic Literature,* 32**,** 1667-1717.

BORJAS, G. & DORAN, K. 2012. The Collapse of the Soviet Union and the Productivity of American Mathematicians. *NBER Working Paper 17800.* Cambridge, Mass.: NBER.

BRESNAHAN, T. & GAMBARDELLA, A. 2004. Old-Economy Inputs for new-Economy Outcomes: What have we learned? *In:* BRESNAHAN, T. & GAMBARDELLA, A. (eds.) *Building High-Tech Clusters.* Cambridge CUP

BRINKLEY, I. 2008. The Knowledge Economy: How knowledge is reshaping the economic life of nations. London: The Work Foundation.

CAMERON, A. C. & TRIVEDI, P. 2009. *Microeconometrics using Stata,* College Station, Stata Press.

CANTWELL, J. 2005. MNCs, local clustering and science-technology relationships. *In:* SANTANGELO, G. (ed.) *Technological Change and Economic Catch-Up: The Role of Science and Multinationals.* Cheltenham: Edward Elgar.

CARD, D. 2005. Is the New Immigration Really So Bad? *The Economic Journal,* 115**,** 300-323.
CASCIANI, D. 2011. *Q&A: UK immigration cap* [Online]. BBC. Available: http://www.bbc.co.uk/news/10436228 [Accessed 14 February 2012].
CHELLARAJ, G., MASKUS, K. & MATTOO, A. 2005. The Contribution of Skilled Immigration and International Graduate Students to U.S. Innovation *World Bank Policy Research Working Paper 3588.* Washington DC: World Bank.
DOCQUIER, F. & RAPOPORT, H. 2011. Globalisation, Brain Drain and Development. *IZA DP 5590.* Bonn: IZA.
EASTERLEY, W. & LEVINE, R. 1997. Africa's Growth Tragedy: Policies and Ethnic Divisions. *Quarterly Journal of Economics,* 112**,** 1203-1250.
FAGERBERG, J. 2005. Innovation: A guide to the literature. *In:* FAGERBERG, J., MOWERY, D. & NELSON, R. (eds.) *The Oxford Handbook of Innovation.* Oxford: OUP.
FAGGIO, G. & OVERMAN, H. 2011. The effect of public sector employment on local labour markets. *Spatial Economics Research Centre Annual Conference.* London.
FAIRLIE, R., KRASHINSKY, H., ZISSIMOPOULOS, J. & KUMAR, K. 2009. Indian Entrepreneurial Success in the United States, Canada and the United Kingdom. *RAND Corporation Working Paper 727.* Santa Monica, CA: RAND Corporation.
FLORIDA, R. 2002. *The Rise of the Creative Class,* New York, Basic Books.
FREEMAN, C. 1987. *Technology Policy and Economic Policy: Lessons from Japan* London, Pinter.
FUJITA, M. & WEBER, S. 2003. Strategic Immigration Policies and Welfare in Heterogenous Countries. *Institute of Economic Research Working Papers.* Kyoto: Kyoto University.
GIBBONS, S., OVERMAN, H. & RESENDE, G. 2011. Real Earnings Disparities in Britain. *SERC Discussion Paper SERCDP0065.* London: LSE.
GORDON, I., WHITEHEAD, C. & TRAVERS, T. 2007. The Impact of Recent Immigration on the London Economy. London: City of London Corporation.
GREENE, W. H. 1994. Accounting for Excess Zeros and Sample Selection in Poisson and Negative Binomial Regression Models. *NYU Working Paper No. EC-94-10.* New York NYU.
HALL, B., JAFFE, A. & TRAJTENBERG, M. 2001. The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools. Cambridge, Mass.: NBER.
HALL, P. 1998. *Cities in Civilisation: Culture, Innovation and Urban Order,* London, Weidenfeld and Nicholson.
HONG, L. & PAGE, S. 2001. Problem Solving by Heterogeneous Agents. *Journal of Economic Theory,* 97**,** 123-163.
HONG, L. & PAGE, S. 2004. Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 16385-16389.
HUNT, J. & GAUTHIER-LOISELLE, M. 2008. How much Does Immigration Boost Innovation? *NBER Working Paper 14312.* Cambridge, Mass.: NBER.
HUNT, J. & GAUTHIER-LOISELLE, M. 2010. How Much Does Immigration Boost Innovation? *American Economic Journal: Macroeconomics,* 2**,** 31-56.
JACOBS, J. 1969. *The Economy of Cities,* London, Vintage.
JAFFE, A. & TRAJTENBERG, M. 1999. International Knowledge Flows: Evidence from patent citations. *International Journal of Knowledge Organisation,* 8**,** 105-136.

JAFFE, A. B., TRAJTENBERG, M. & HENDERSON, R. 1993. Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *The Quarterly Journal of Economics,* 108**,** 577-598.

KAISER, U., KONGSTED, H. C. & RØNDE, T. 2011. Labor Mobility, Social Network Effects, and Innovative Activity. *In:* 5654, I. D. N. (ed.). Bonn: IZA.

KAPUR, D. & MCHALE, J. 2005. Sojourns and Software: Internationally mobile human capital and high tech industry development in India, Ireland and Israel. *In:* ARORA, A. & GAMBARDELLA, A. (eds.) *From Underdogs to Tigers: The Rise and Growth of the Software Industry in Brazil, China, India, Ireland and Israel.* Oxford: OUP.

KERR, S. P. & KERR, W. 2011. Economic Impacts of Immigration: A Survey *NBER Working Paper 16736.* Cambridge, MA: NBER

KERR, W. 2008a. The Agglomeration of US Ethnic Inventors. *HBS Working Paper 09-003.* Boston, MA: Harvard Business School.

KERR, W. 2008b. Ethnic Scientific Communities and International Technology Diffusion. *Review of Economics and Statistics,* 90**,** 518-537.

KERR, W. 2009. Breakthrough Innovations and Migrating Clusters of Innovation. *NBER Working Paper 15443.* Cambridge, MA: NBER.

KERR, W. & LINCOLN, W. 2010. The Supply Side of Innovation: H-1b Visa Reforms and US Ethnic Invention *NBER Working Paper 15768.* Cambridge, Mass.: NBER

KYAMBI, S. 2005. *Beyond Black and White: Mapping New Immigrant Communities,* London, ippr.

LAKHA, F., GORMAN, D. & MATEOS, P. 2011. Name analysis to classify populations by ethnicity in public health: Validation of Onomap in Scotland. *Public Health,* 125**,** 688-96.

LANDRY, C. & WOOD, P. 2008. *The Intercultural City: Planning for Diversity Advantage,* London, Earthscan.

LEADBEATER, C. 2008. The Difference Dividend: Why Immigration is Vital to Innovation. *NESTA Provocation.* London: NESTA.

LEGRAIN, P. 2006. *Immigrants: Your Country Needs Them,* London, Little, Brown.

LI, X. & PAI, Y. 2010. The Changing Geography of Innovation Activities: What do Patent Indicators Imply? *In:* FU, X. & SOETE, L. (eds.) *The Rise of Technological Power in the South.* Basingstoke: Palgrave MacMillan.

LISSONI, F., TARASCONI, G. & SANDITOV, B. 2006. The KEINS Database on Academic Inventors: Methodology and Contents. *CESPRI Working Paper 181.* Milan: Universita' Bocconi.

MANNING, A. & ROY, S. 2007. Culture Clash or Culture Club? The Identity and Attitudes of Immigrants in Britain. *CEP Discussion Paper 790.* London: Centre for Economic Performance, London School of Economics.

MARÉ, D. C., FABLING, R. & STILLMAN, S. 2011. Immigration and Innovation. *IZA Discussion Paper 5686.* Bonn: IZA.

MATEOS, P., WEBBER, R. & LONGLEY, P. 2007. The Cultural, Ethnic and Linguistic Classification of Populations and Neighbourhoods using Personal Names. *CASA Working Paper.* London: UCL.

MENON, C. 2009. Stars and Comets: An Exploration of the Patent Universe. *SERC DP0037.* London: LSE.

MOWERY, D. C. 2001. Technological Innovation in a Multipolar System: Analysis and Implications for U.S. Policy. *Technological Forecasting and Social Change,* 67**,** 143-157.

NAKHAIE, R., LIN, X. & GUEN, G. 2009. Social Capital and the Myth of Minority Self-Employment: Evidence from Canada. *Journal of Ethnic and Migration Studies,* 35**,** 625-644.
NATHAN, M. & LEE, N. 2011. Does Cultural Diversity Help Innovation in Cities? Evidence from London Firms. *Spatial Economics Research Centre Discussion Paper SERCDP0069.* London: LSE.
NIEBUHR, A. 2006. Migration and Innovation: Does Cultural Diversity Matter for Regional R&D Activity? *IAB Discussion Paper 14/2006.* Nuernberg: Institut fuer Arbeitsmarkt- und Berufsforschung.
OECD 2009. OECD Patent Statistics Manual. Paris: OECD.
OTTAVIANO, G., BELLINI, E. & MAGLIETTA, A. 2007. Diversity and the Creative Capacity of Cities and Regions. *SUSDIV Paper 2.2007.* Bologna: FEEM.
OTTAVIANO, G. & PERI, G. 2005. Cities and Cultures. *Journal of Urban Economics,* 58**,** 304-337.
OTTAVIANO, G. & PERI, G. 2006. The Economic Value of Cultural Diversity: Evidence from US Cities. *Journal of Economic Geography,* 6**,** 9-44.
OZGEN, C., NIJKAMP, P. & POOT, J. 2010. Immigration and Innovation in European Regions. *Department of Spatial Economics working paper.* Amsterdam: VU University.
OZGEN, C., NIJKAMP, P. & POOT, J. 2011. The Impact of Foreign Workers on Innovation and Knowledge Spillovers: Analysis of Survey Data from the Netherlands. *Migration: Economic Change*, *Social Challenge.* London: UCL.
PAGE, S. 2007. *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools and Societies,* Princeton, Princeton University Press.
PARROTTA, P., POZZOLI, D. & PYTLIKOVA, M. 2011. The Nexus Between Labor Diversity and Firm's Innovation [sic]. *NORFACE Discussion Paper 2011-5.*
PERI, G. 2007. Higher Education, Innovation and Growth. *In:* BRUNELLO, G., GARIBALDI, P. & WASMER, E. (eds.) *Education and Training in Europe.* Oxford: Oxford University Press.
ROBSON, B., BARR, R., LYMPEROPOULOU, K., REES, J. & COOMBES, M. 2006. A framework for City-Regions: Working Paper 1: Mapping City-Regions. London: Office of the Deputy Prime Minister.
RODRÍGUEZ-POSE, A. & STORPER, M. 2006. Better Rules or Stronger Communities? On the social foundations of institutional change and its economic effects. *Economic Geography,* 82**,** 1-25.
ROMER, P. 1990. Endogenous Technological Change. *Journal of Political Economy,* 98**,** 71-102.
SAXENIAN, A.-L. 2006. *The New Argonauts: Regional Advantage in a Global Economy,* Cambridge, MA, Harvard University Press.
SAXENIAN, A.-L. & SABEL, C. 2008. Venture Capital in the 'Periphery': The New Argonauts, Global Search and Local Institution-Building. *Economic Geography,* 84**,** 379-394.
SCHUMPETER, J. 1962. *The Theory of Economic Development*, Berlin, Springer.
SIMMIE, J., CARPENTER, J., CHADWICK, A. & MARTIN, R. 2008. *History Matters: Path dependence and innovation in British city-regions,* London, NESTA.
STEPHAN, P. & LEVIN, S. 2001. Exceptional Contributions to US Science by the foreign-born and foreign-educated. *Population Research and Policy Review,* 20**,** 59-79.
WADHWA, V., SAXENIAN, A.-L., RISSING, B. & GEREFFI, G. 2007. America's New Immigrant Entrepreneurs. Durham, NC: Duke University / iSchool, UC Berkeley.

WADSWORTH, J. 2010. Immigration and the UK Labour Market: The Evidence from Economic Research. *CEP Election Analysis.* London: LSE.
WILSON, R. 2007. Innovation in London. *Working Paper 19.* London: GLA Economics.
YEUNG, H. 2009. Regional Development and the Competitive Dynamics of Global Production Networks: An East Asian Perspective. *Regional Studies,* 43**,** 325-351.
ZENOU, Y. 2011. Spatial versus Social Mismatch: The Strength of Weak Ties. *IZA Discussion Papers 5507.* Bonn: Institute for the Study of Labor (IZA).