




instituteofculturalcapital

The Culture and Education District
A scoping review for the Arts and Humanities Research Council

Directed by Dr Beatriz Garcia, ICC

Appendix B: Economic Work Strand Report

Max Nathan
Henry Overman
Sevrin Waights

What Works Centre for Local Economic Growth
February 2018



Contents

1. Aims and objectives	3
2. Methods	4
3. Literature review	6
3.1. Culture	6
3.2. Education	7
3.3. Physical regeneration	8
3.4. Area-based incentives and support for firms	10
3.5. Impact evaluation challenges and appropriate methods	11
3.5.1. Challenges for ex-post impact evaluation	11
3.5.2. Appropriate methods	13
4. Datasets review	14
4.1. Firm data.....	14
4.2. Household/individual datasets	15
4.3. Area-level datasets	15
4.4. 'Big data'	16
5. Recommendations	18
6. References	19
7. Appendices	21

1. Aims and objectives

The potential economic impacts of the Culture and Education District (CED) are diverse and wide ranging. If we view the CED as creating density or new networks (or both) then there are many channels through which this could potentially impact the local economy, as well as the wider city-region. A diverse array of disciplines have considered these effects using many different methodologies. In economics alone, recent handbook and review papers summarising the evidence on agglomeration economies (Combes and Gobillon 2015; Duranton et al. 2015) and networks (Jackson, 2008; Jackson 2005) each provide references to several hundred papers.

To the extent that the CED provides new data and new case studies, the possibilities for additional work on economic effects and impact channels – that builds and extends the existing literature – is considerable, and it makes little sense to try to summarise it here. Instead, we take a narrower approach and consider what we might learn from the CED viewed as a policy intervention (whose objective is to facilitate increased clustering, and through this, to improve some set of economic outcomes in the surrounding neighbourhoods and for London).

Even within this narrower framing, there is an extensive literature that seeks to understand the development and the effects of cluster policy, both at area level and on specific sets of firms, workers and residents in and around the ‘treated’ locations. This literature considers a wide-ranging set of economic (and non-economic) outcomes, and uses an array of research designs. As with the broader literature on density and networks it makes little sense to try to summarise it here other than to note that CED will provide plenty of opportunities for work along these lines.

Instead, we focus on evaluations that identify effects which can be attributed, with some degree of certainty, to the support provided – and the scope for the CED to allow for additional evaluation work that will help us better understand those economic impacts.

2. Methods

Specifically, this section looks at the potential economic impacts of the Culture and Education District (CED) by drawing on findings from the What Works Centre for Local Economic Growth (hereafter WWC). In particular, we draw on large-scale systematic reviews and smaller-scale toolkits, as well as a review of relevant UK datasets. The aim is to map and review the relevant evidence base for CED-type interventions and associated policies; to identify some of the challenges in assessing impacts and propose appropriate research design strategies; and to identify UK datasets that could be used in such research.

The WWC focuses on specific kinds of evidence, and it is important to be clear what we are including.¹ First, the WWC reviews impact evaluations that look to identify the effect of some policy on some outcome, rather than forecasts or projections of impact (which could turn out to be wrong). These evidence reviews draw on studies from across the OECD, with no time limit on study date.

Second, the WWC uses a standards of evidence filter to assess studies that are robust enough (in terms of research design and execution) to give us confidence that any impacts are real. The filter is based on the Scientific Maryland Scale, which runs from SMS5 (a Randomised Control Trial) to SMS1 (basic cross section). The minimum standard of any study cited below is SMS2: that is, a before/after analysis or cross-section analysis in which controls have been fitted to try and isolate the policy effect from confounding factors.

Many CED-type policies involve large amounts of capital spending, for example on physical infrastructure: in these cases Randomised Control Trials (RCTs) are almost always not suitable, given the single site, long time frame, and difficulty in reversing any negative effects once these emerge. In these cases other 'quasi-experimental' methods such as Synthetic Controls, or observational methods using decent quality data sets, will be preferred for impact evaluation. However, for programmes complementary to the CED, such as active labour market or business support initiatives, randomisation may well be an option in evaluation design. We discuss these issues in more detail in Sections 3.5 and 4.

The CED is a 'mega-project' with many component parts, and where the objectives and means to achieve these objectives are still being developed. This latter point means that a precise evaluation of likely impacts cannot be done at this stage. However, we can usefully review the evidence on CED-type interventions, both at 'mega-project' scale and looking at the likely components. We therefore proceed as follows:

¹ For more detail on the WWC approach, see www.whatworksgrowth.org.

- We review the evidence on the effect of cultural events and facilities to discuss the potential economic impact of the cultural interventions proposed in the CED (e.g. the two art galleries, the theatre, etc.)
- In order to assess the potential impact of the educational interventions (in particular the two university campuses), the section draws on two policy toolkits on the impact of university incubators and co-location, and on a study of the impact of public sector relocation.
- We summarise relevant evidence from two reviews of physical regeneration (estate renewal and public realm interventions), as well as a review on area-based initiatives that provide incentives or support to businesses (such as enterprise zones).

Finally, two broader points are worth making before we go into the detail. First, interventions such as these are often used with the aim of neighbourhood regeneration. Specifically, neighbourhoods with high rates of deprivation often suffer from poor access to arts, culture, educational facilities, and disinvestment in the built environment. The hope is that by investing in such improvements, an initiative can improve the life chances of local residents. This means that a key issue in evaluating policy impacts is looking at **who gains or loses out from the policy** (at individual, household or firm level), not just at **what changes on the aggregate** (usually at area level). The following presents an overview of the international empirical evidence of the economic effects of similar interventions.

A second big point for evaluators is that ideally, we want to look at the economic (and non-economic) impacts of both CED **component parts and projects**, and that of **the CED as a whole**. The latter, in particular, may be both long-term and hard to see (especially as component parts and projects may interact with each other). We return to some of these evaluation challenges below.

3. Literature review

3.1. Culture

Summary: There is surprisingly little robust evidence on the local economic impacts of cultural facilities and events (compared to sports facilities and events, for example). On the very limited evidence we have – three studies – the key findings are: cultural interventions do not always impact positively on the local economy; and they may improve outcomes via gentrification rather than benefiting existing residents; that is, areas ‘improve’ but existing actors don’t gain (and may be displaced). Most pertinently, a major US study on cultural districts finds both positive aggregate economic impacts, and evidence of displacement. Note that we focus only on economic impacts here, not intrinsic benefits. Also, note that we have not found any direct comparators to the CED in terms of scale or number of actors.

There is surprisingly little robust evidence on the local economic impacts of cultural facilities and events (compared to sports facilities and events, for example, where many high quality studies exist).² On the very limited evidence we have, the key findings are:

- Cultural interventions do not always impact positively on the local economy
- They may improve outcomes via gentrification rather than benefiting existing residents; that is, areas ‘improve’ but existing actors don’t gain from this (and may be displaced).

Note that we focus only on *economic* impacts here, not any intrinsic benefits. Also, we have not found any direct comparators to the CED in terms of scale or number of actors. However, the studies reviewed here are relevant to many potential components of the CED proposals.

These findings chime with the broader evidence base on clusters and how they develop (see above). However, we would caution that here we have a case of ‘little evidence of impact’ rather than ‘evidence of little impact’. For that reason, further high quality evaluation is needed. The CED programme presents a good opportunity to do this.

Specifically, in the WWC *Evidence Review on Sports & Culture*, only three robust evaluations were found on the impact of cultural projects. These studies are worth examining individually since they all contain aspects of arts-led regeneration that is relevant to the CED. The first study examines 99 Cultural Districts in the U.S. over the 1990s and 2000s, and is perhaps the most relevant to the CED programme. The study finds that cultural

² The WWC review of sports and culture covers the sports evidence. See <http://www.whatworksgrowth.org/policy-reviews/sports-and-culture/>

districts have a significant positive effect on local property prices, employment and income, both within the Districts and in adjacent neighbourhoods. But it also finds an increase in the share of college educated residents, a decrease in the number of residents living in poverty, and an increase in residential turnover. These last effects, and the increase in local property prices are indicative of gentrification and potentially displacement.

The second study examines the impact of the location of art galleries in New York City on neighbourhood redevelopment – measured by building stock and land use changes over time. The study finds that galleries do not lead to redevelopment.

The third study examines the effect on residents' quality of life and local GDP per capita of hosting the European Capital of Culture for 21 cities between 1994-2005. The study finds no effect on city-level GDP per capita and a negative effect on quality of life in the hosting year, which returns to previous levels the year after. The authors suggest the negative effect may be the result of dissatisfaction with transport disruptions or general overcrowding.

3.2. Education

Summary: We draw on a wider range of evidence bases to explore potential economic impacts of the CED's 'education' elements. Reviewing 14 studies, the key findings are: universities may help local firms to bring research to market (commercialisation); university business incubators may improve outcomes (e.g. employees) for supported firms; but relocation of university jobs (at the scale of a few hundred staff) is not likely to create many additional private sector jobs. Again, note that we have not found any direct comparators to the CED in existing relocation studies; the CED is more comparable in scale to the BBC's move to MediaCity than any recent UK government relocation exercises.

We draw on a wider range of evidence bases to explore potential economic impacts of the 'education' elements of the CED. Specifically, we are able to look at university research commercialization activities, both through Technology Transfer Offices and more broadly; incubator spaces (and accelerator programmes); and small-scale public sector relocation.³

On the evidence we have, the key findings are:

- Universities may help local firms bring research to market (commercialisation)
- University business incubators may improve outcomes (e.g. employees) for supported firms

³ We follow the definition of 'incubators' in Hathaway (2016). We define incubators as distinct from science parks / business parks that co-locate firms but do not actively mentor/encourage collaboration.

- Relocation of university jobs at small scale is not likely to create additional private sector jobs.

Note that the CED is more comparable in scale to the BBC's move to MediaCity than any recent UK government relocation exercises. The WWC is developing an internal project to look at the labour market impacts of the MediaCity move; early results should be ready before end-2017.

Specifically, we draw on evidence from the WWC *Toolkit on Colocation for Commercialisation* (4 studies). This evidence suggests that the process of commercialising research (measured by patents) benefits from geographical proximity. Further, firms are likely to locate their R&D branches within 10km of universities with relevant research departments. These findings imply that the two universities in the CED may provide benefits to local firms or even attract new R&D establishments.

We also draw on our *Toolkit on Business Incubators* (8 studies). This evidence suggests that incubators may have positive economic effects for supported firms in terms of employees and sales. This means that the CED has the potential to support local firms if the universities engage in business incubation. However, it is unclear the extent to which this may contribute to outcomes at the area level.

Finally, we have a couple of studies on public sector relocation. The WWC has yet to conduct a systematic review of the evidence in this area. There are no robust evaluations of large scale one-off UK public sector relocations such as BBC Media City. However, there are two important studies of the impact of changes in public sector jobs on local economies. One study (Faggio & Overman 2014) examines all public sector job changes over 2003-2007, most of which were due to expansion of healthcare, finding that the net effect on private sector jobs is zero. The other study (Faggio 2015) examines specific relocations of 25,000 public sector jobs around the country, finding that 10 public sector jobs create 5.5 private sector jobs and that the impact on jobs occurs within 2km of the relocation sites as a result of the Lyons Review. Both studies find that relocation of public sector jobs increases services employment at the expense of manufacturing jobs.

3.3. Physical regeneration

Summary: We look at two pertinent aspects of physical development: public realm improvements, and improvements to housing and housing estates. Reviewing 21 studies, the key findings are: physical improvements (public realm and estate renewal) are not likely to directly impact on local employment or productivity; however, associated policies (e.g. active labour market programmes) may help tie local residents into work; physical improvement programmes may be associated with localised increases in house prices and

rents; such programmes often explicitly aim to attract new residents; it is unclear that existing residents always benefit from this: house price/rent increases and/or development models can induce displacement.

We look at two pertinent aspects of physical development: public realm improvements, and improvements to housing and housing estates. Perhaps not surprisingly, to date there has been no impact evaluation of public realm interventions. Especially at scale, these are complex interventions where impacts are hard to track. However, it is possible to develop a framework for thinking-through impacts, and we summarise one of these, for economic outcomes, below. There is a rather better evidence base for estate renewal, with 21 evaluations in the WWC systematic review. On this framework and evidence base, the key findings are:

- Physical improvements (public realm and estate renewal) are not likely to directly impact on local employment or productivity; although associated policies (e.g. active labour market programmes) may help tie local residents into work
- Physical improvement programmes may be associated with localised increases in house prices and rents
- Such programmes often attract new residents, and this is often an explicit aim of the policy; it is unclear that existing residents always benefit, to the extent that house price/rent increases and/or development models induce displacement.

We draw first on the WWC *Review of Public Realm Improvements*. A systematic review did not yield any sufficiently robust evaluations on the economic impacts of such interventions. Nevertheless the WWC made use of related research to produce a policy briefing on the potential economic effects of such interventions. The briefing came to the following conclusions. Public realm interventions may help create more attractive places to live, but this might lead to higher housing costs and displace existing residents. Large-scale public realm interventions may help attract new residents and create mixed communities but it is unclear that simply living in mixed communities delivers significant economic benefits to existing residents. Smaller-scale public realm interventions that aim to modestly improve the wellbeing of existing residents are likely to have fewer undesirable or unintended consequences than large-scale, radical transformations. Public realm improvements in commercial areas might boost overall business activity but will not necessarily increase jobs or firm profits in the long term, and might also displace existing businesses.

As noted above, the WWC also conducted a systematic review on *Estate Renewal Programmes* (21 evaluations). Estate renewal programmes lead to increases in property and land prices and rents, although not always for nearby properties that do not directly benefit from improvements. Estate renewal programmes tend to have a limited impact on the local area in terms of improving poverty or employment rates. There is some evidence that estate renewal increases income at the area level; however, the studies are unable to

distinguish the extent to which these results are driven by improvements for existing residents or changes in neighbourhood composition. Estate renewal programmes tend to have a limited impact on the local area in terms of reducing area-level crime, improving health, wellbeing or education. Evidence of impacts on existing residents is often not measured in these studies.

3.4. Area-based incentives and support for firms

Summary: Typically such programmes involve identifying a set of ‘zones’ for growth, and applying a combination of land/buildings remediation, infrastructure improvements and/or business incentives (such as tax breaks) for firms who locate in these areas. Reviewing 30 evaluations, the key findings are: just over 50% of studies report positive effects on area-level labour market outcomes; around half report positive effects on area-level poverty or wages; a number of studies suggests that these positive area-level effects for area-based initiatives can be driven by displacement from other areas; programmes may work best for local residents when local hiring conditions are used.

In this section we look at a range of area-based initiatives (ABIs) which aim to raise local economic growth. Typically these programmes involve identifying a set of ‘zones’ for growth, and applying a combination of land/buildings remediation, infrastructure improvements and/or business incentives (such as tax breaks) for firms who locate in these areas. A key issue with this programme design, then, is the extent to which these incentives simply shift activity from one part of a city to another, rather than generating net new activity. A related issue is that if existing firms move into the zone, this may not result in new jobs for local people. In some cases, therefore – such as the US Empowerment Zone model – additional benefits are received when firms hire local residents.

It is not clear, at the moment, whether these kind of programmes will form part of the CED, but it is safe to assume they could be part of the policy mix. We are able to draw on a review of 30 studies evaluating economic ABIs. On this evidence base, the key findings are:

- Just over 50% of studies report positive effects on area-level labour market outcomes; around half report positive effects on area-level poverty or wages
- A number of studies suggests that these positive area-level effects for area-based initiatives can be driven by displacement from other areas
- Economic ABIs may work best for local residents when local hiring conditions are used.

Specifically, we draw on the *WWC Evidence Review on Enterprise Zones and Empowerment Zones and other ABIs* (30 studies). This finds that a little over half the reviews find positive impacts on zone-level employment/unemployment. Half of the studies that

consider the impact on poverty or wages report positive effects. Most of the reviews that consider the number of businesses find positive effects.

A number of studies suggest that positive effects for Zones may be driven by displacement from nearby areas. For example one study of the UK's Local Enterprise Growth Initiative (LEGI) found the policy led to increases in employment inside treated areas, but at the expense of decreases in employment in nearby untreated areas.

One study of US Enterprise Zones makes use of variation in programme design across states to suggest that employment growth in existing target businesses is promoted only if programme incentives are tied to local hiring requirements. In particular, for the EZs considered, employment growth was an outcome only where tax incentives were tied to the creation of jobs. U.S. Empowerment Zones, examined in another study take this a step further and provide firms with hiring tax credits only for local residents. It could be a result of these conditions that the study finds that Empowerment Zones perform better at increasing local employment than Enterprise Zones.

3.5. Impact evaluation challenges and appropriate methods

Evaluating the economic impacts of the CED will be challenging. As noted above, it is a mega-project with various component parts; and where specific objectives and policies to achieve these are under development. It is also likely that while some CED components will achieve results (or not) within a year or so, the effect of others will not be known for some time after. It is also plausible that individual components of the CED will interact with each other, so that the aggregate impact of the intervention may be greater than the sum of its parts (or less, if the effects of one element work against another).

In what follows, we look at the major evaluation challenges involved in testing the economic impacts of the CED and CED components; we then review some of the quantitative methods that can be used by policymakers and researchers.

3.5.1. Challenges for ex-post impact evaluation

The major challenge to evaluating the economic impact of any policy intervention is finding a suitable 'counterfactual', i.e. to know would have happened without the support. It is not enough to know what happens to an area after an intervention – you also need to know how much of the change is attributable to the policy and how much would have happened in any case. Estimating a counterfactual is made all the more difficult by the fact that supported areas tend to be 'selected' into a policy treatment. For a programme like the CED, there are three main challenges.

Location

First, the CED location has distinctive characteristics and history that make it unlike other parts of London (and arguably, unlike anywhere else in the UK). Similarly, the levels of investment and policy attention that have gone into the area since 2005 are also unusual, perhaps unique. In economic language, the area is 'selected', both on past economic and social trends and on growth potential. This means that CED neighbourhoods are likely to experience different economic trends from other parts of London and from other cities. While some of these trends will be explicable in terms of observable characteristics, and so can be controlled for in analysis, others will be harder to observe but may still shape outcomes.

The selection problem is very significant for the CED. The CED is based in the Borough of Newham, which is an area with high rates of deprivation. The Olympic site itself is the focus of a wide range of initiatives such as the Here East tech cluster, transport improvements, and the new usage of the Olympic stadium itself as a football stadium. Overall, the area targeted by the CED is likely to experience different economic trends to other areas for reasons that may have nothing to do with the CED itself. This means evaluation of the CED presents a challenge of how to disentangle the effects of the CED from other factors impacting on the area.

Displacement effects

Second, as we have seen, the type of interventions involved in the CED tend to be associated with displacement effects. Residents may move away if housing costs increase. This adds a layer of complexity to any policy evaluation since, for example, increased area level wages may be due to a positive effect on wages or may result because low income households have moved away to be replaced by higher income households. Specifically, it means that evaluators will need high quality data that is ideally able to follow individuals / households and firms over time. This is crucial if we are to explore what economists call the 'distributional effects' of the CED; that is, the winners and losers from the policy.

Single site

Third, the CED involves only a single site, and as we have seen, has many unusual / unique characteristics. This is problematic for an evaluation examining the impact on area-level outcomes. The CED will only impact on very few geographical areas in its immediate vicinity, depending on the size of the units used. This makes it difficult to construct a treatment and control group of the necessary size to evaluate the effect of a policy.

3.5.2. Appropriate methods

In order to deal with selection problems many evaluations use variations on difference-in-difference or panel fixed effects methods.⁴ In these methods, the change in outcome in a 'treatment' group (those that receive support) is compared with the change in outcome in a group of similar control areas (which do not). Control group units are carefully selected to be as similar as possible to treated units, sometimes using matching techniques and sometimes adjusting for differences using control variables.

Successful implementation of the diff-in-diff approach would also make use of differences in the timing and geography of interventions to isolate the effect of the CED. For example, one might use firm data with precise address location, to see how changes in employment are associated with distance to the CED whilst controlling for distance to other interventions such as the Here East tech cluster. This is only possible because the Here East tech cluster is not in the exact same location as the CED – rather it is about 1km away from the CED in the direction of Hackney. On the other hand, the London Stadium and the London aquatic centre are very close to the CED, so disentangling would be based on timing. This could work by looking at changes in outcomes specifically focussed around the implementation date of the CED, controlling for broader changes that may be driven by the renovation of the London Stadium over the period 2013-2016.

In order to deal with potential residential displacement effects, an evaluation should track individuals that moved away from the area. The effect on original residents can then be estimated as the change in outcomes vs the control group for individuals that stayed and those that moved away. Another form of displacement is if firms locate inside the CED because of incentives or the benefits of being close to a university instead of locating elsewhere in Stratford. This effect could be estimated by examining area-level changes both inside the CED and outside the CED compared with a control group. If the CED simply displaces firms then a positive effect inside would be met with a negative effect in areas outside of equal magnitude.

Finally, the problem of the CED targeting only a single site means that the method should use small geographical units, or individual firm/household data in order to construct treatment and control groups of sufficient sample size to estimate an effect. Another possibility would be to make use of emerging methods where one treated area is compared to a 'synthetic' control area that is constructed based on aggregation of many potential control sites.

⁴ For more information on the methods referred to in this section, see <http://www.whatworksgrowth.org/resources/scoring-guide/>. Synthetic control methods are discussed here with reference to the BBC's move to Manchester: <http://www.whatworksgrowth.org/blog/location-location/>

4. Datasets review

This section reviews the main datasets that could be used in analysis of the economic impacts of the Cultural and Education District. We start with firm level data, then look at individual/household data, area-level datasets and 'big data' resources. Appendix Table 1 summarises key economic variables that researchers might be interested in, and datasets available for measuring each. For each dataset, Appendix Table 2 provides information on access arrangements and cost, if any.

4.1. Firm data

The **Business Structure Database (BSD)** gathers company information through HMRC tax records. It contains basic information such as firm birth/death year, employee count, SIC code, and turnover for the entire population of around 4 million firms and 5.5 million local entities (plants) in the UK in any year. Its large number of observations and the availability of firm British National Grid coordinates makes the BSD the preferred firm dataset for the variables that it covers, especially for analysis at a fine geographical scale. An alternative to BSD is **Companies House (CH)** data, which contains information on around 3 million companies registered in the UK. The CH data contains more information than BSD – such as assets and profits and losses but only electronically filed company accounts are available (around 75% of the total). An advantage of CH data is that it is more likely to include very new start-ups than the BSD, since the latter requires the business to be integrated in the tax system. This means CH it is often preferred for examining business formation. There are a number of disadvantages to the CH that mean the BSD would usually be preferred for other outcomes. One such disadvantage is that the CH is a non-random sample of firms that filed electronically. Another is that the registration address is not necessarily a location of operations. Finally, since the data are accessed via a web API, there may be some programming challenges and fees for unrestricted access.

The Experian **National Business Database (NBD)** is a further alternative that is based on CH data combined with entries from the Yellow Pages and Thompson business directories. Combining these sources provides a larger set of firms overall; however, there is likely to be only limited information for those firms found exclusively in Yellow Pages. Further, this dataset is currently only available for three years: 2013, 2014 and 2015. Given the excellent coverage, it is likely to be a good dataset for examining firm formation.

The **Annual Business Survey (ABS)** and the **Annual Respondents Database X (ARDx)** have much smaller sample sizes – only 50,000 firms per year (approximately 1% samples) – but collect more information on firms. For example, the ABS has information on labour and other costs of production. The ARDx is a sub-sample of the BSD and may be used to complement it with extra information. It has information on capital stock and gross value

added, as well as employment, which together would allow for estimation of total factor productivity. However, the sample sizes for ABS and ARDx imply only about 15 firms for the Stratford and New Town ward (if roughly proportional with its population of about 20,000), meaning it is unlikely that a fine scale analysis would be possible with these datasets.⁵ Further, the ARDx does not cover firms from all sectors. The **Business Register and Employment Survey (BRES)** has around 80,000 firms and 500,000 local entities. As with BSD, it is based on tax records. Since it doesn't include much more information than the BSD, the BSD would usually be preferred. The **Business Expenditure on Research and Development (BERD)** contains information on firm-level research and development expenditure; however, since it has a fairly small sample of around 30,000 firms, it is not suitable for a fine scale analysis.

4.2. Household/individual datasets

Both the **Annual Population Survey (APS)** and the **Quarterly Labour Force Survey (QLFS)** each contain around 500,000 individuals per year and have locations to the Output Area level. This sample size implies around 120 individuals in the Stratford and New Town ward if proportional with population. By pooling several pre- and post-treatment years together, a sufficient sample for a fine scale analysis could potentially be reached. A fine scale analysis could potentially pool both datasets to boost the sample size for variables covered by both datasets, such as economic status. It is unlikely that the sample size would be sufficient for an area level analysis (e.g. impacts on share of residents holding a university degree). The APS contains information on subjective well-being, and the QLFS contains data on income.

The **Annual Survey of Hours and Earnings (ASHE)** has a smaller sample size of around 150,000 to the postcode level. Whilst the sample is smaller, it is a longitudinal dataset and may be used for a more robust analysis compared with the QLFS, which only has a panel element, or the APS, which is a repeated cross-section. Further, the panel element makes it the best option for examining residential mobility. Again, by pooling several years before and after the treatment, it might be possible to build a sample of a few hundred observations in the vicinity of the treatment. The ASHE contains data on wages and hours of work.

4.3. Area-level datasets

The **UK Census** could potentially be used to examine impact for a variety of outcomes. In particular it may be the only dataset to reliably analyse demographic change (e.g. degree share, ethnic group) or total population for areas smaller than wards or MSOAs. A major

⁵ This is not to say that the analysis should be confined to this ward, but to give a rough indication of the number of observations in the vicinity of the Queen Elizabeth Olympic Park.

drawback is that it is only released every ten years. However, changes between 2011 and 2021 could be used to evaluate the effects of policies implemented in that period with the change between 2001 and 2011 being used to control for pre-trends.

There are **mid-year population estimates** available at the LSOA level from the Greater London Authority Datastore but it is unlikely that these model-based figures are reliable for evaluation purposes. The **Land Registry** (LR) provides transaction Price Paid Data for residential housing unit sales. The data includes the entire population of sales since 1995 (more than 24 million records). The **Recorded Crime Survey (RCS)** provide data on crimes down to the LSOA level. The **Index of Multiple Deprivation (IMD)** provides a measure of deprivation at a fine local level, but it is only available every few years.

The Business and Local Government Data Research Centre (BLGDRC) holds datasets on **outstanding SME loans** and **outstanding personal loans** at the area-level for postcode sectors. These datasets are available on a quarterly basis since 2013.

4.4. 'Big data'

A lack of conventional data may be addressed using 'big data' sources. For example, there is a lack of detailed income data in the UK. In the US, **Google Street View (GSV)** has been used previously to predict neighbourhood income. One study does this by using an algorithm to detect the make and model of cars parked along streets, and matching this to information about the car including its price (Gebru et al. 2017). Another study does this by processing the image in terms of the built structures (i.e. colours, textures, etc.) to predict income of the residents (Glaeser et al. 2016). However, both of these studies use other sources of income data to 'calibrate' the predictions – something that may be difficult in the UK. Nevertheless, it is an approach to bear in mind.

Property listings online, such as Zoopla, may provide an alternative source of house price data to the Land Registry data. Further, they may be the only possible source for rental prices, and for commercial and land transactions. Land transactions, for example, may be found on the UK Land Directory. Commercial listings may be found at Rightmove.

Neighbourhood redevelopment is an outcome which lacks good traditional data. There is fairly poor information on dwelling stock count and condition at the local level, as well as land use. Google Street View, as mentioned above, could be used for this purpose. Further, analysis using geo-tagged tweets from **Twitter** or photos from sites such as **Flickr** can reveal what resident think about certain neighbourhoods in terms of aesthetics or other factors (Ahlfedlt 2013).

The Consumer Data Centre (CDRC) holds a variety of **GeoData packs** outcomes of potential interest. For example it has GPS data on cyclist movements from users of the 'Bike Citizens' app in London. This could be useful for evaluating the impacts of public realm improvement that include cycleways. Other GeoData packs are available for things such as broadband network infrastructure, energy consumption, commuting patterns and WiFi hotspots.

5. Recommendations

Considered as a policy intervention, mega-projects such as the CED are likely to have a range of direct and indirect economic impacts. The evidence base on the effects of CED-type interventions and their component parts is limited, and the CED has arguably unique features. Some of the aggregate impacts of the CED may be long term and hard to track. The AHRC should work with local policymakers to get a better sense of CED aims, objectives and delivery tools.

The available evidence suggests the potential for a range of positive economic outcomes from the CED, both at the aggregate level and at the level of individual firms / residents / households. It also suggests the need for caution about both the extent of economic impacts and their distributional effects. Specifically, there are risks that local residents and/or firms will not benefit, and may active lose out (e.g. through gentrification / displacement). The AHRC should communicate these findings to stakeholders.

There are now a number of quantitative research methods that can be used to robustly evaluate these economic (and non-economic) impacts. The AHRC should directly commission impact evaluation research on the economic effects of the CED and its component parts, and on their distributional impacts. Quantitative projects should be required to meet minimum standards in research design – such as those used by the What Works Centre for Local Economic Growth – as a condition of funding. We encourage the AHRC to commission process evaluations, using qualitative methods, as a key complement to quantitative impact evaluations.

6. References

Ahlfeldt, Gabriel M., Urbanity (2013). CESifo Working Paper Series No. 4533. Available at SSRN: <https://ssrn.com/abstract=2378261>.

Combes P-P. and Gobillon L. (2015) 'The Empirics of Agglomeration Economies.' In: Duranton G., Henderson J. V. and Strange W. (eds) *Handbook of Regional and Urban Economics: Volume 5*. North Holland: Elsevier.

Duranton G., Henderson J. V. and Strange W. (2015) *Handbook of Regional and Urban Economics: Volume 5*. North Holland: Elsevier.

Faggio G. (2015) 'Relocation of Public Sector Workers: The local labour market impact of the Lyons Review.' SERC Discussion Paper 155. London: LSE.

Faggio G. and Overman H. (2014) 'The effect of public sector employment on local labour markets.' *Journal of Urban Economics* 79: 91-107.

Gebu, T., Krause, J., Wang, Y., Chen, D., Deng, J., Aiden, E. L., & Fei-Fei, L. (2017). 'Using deep learning and google street view to estimate the demographic makeup of the US.' arXiv preprint arXiv:1702.06683

Glaeser, E. L., Kominers, S. D., Luca, M., & Naik, N. (2016). 'Big data and big cities: The promises and limitations of improved measures of urban life.' *Economic Inquiry*.

Hathaway I. (2016) 'What Startup Accelerators Really Do.' *Harvard Business Review*, March.

Jackson M. (2008) *Social and Economic Networks*. Princeton: Princeton University Press.

Jackson M. (2005) 'The economics of social networks.' Social Science Working Paper 1237. Pasadena: California Institute of Technology.

What Works Centre for Local Economic Growth. (forthcoming) *Toolkit: Incubator programmes*. London: WWCLEG.

What Works Centre for Local Economic Growth. (2017) *Toolkit: Co-location for Commercialisation*. London: WWCLEG.

What Works Centre for Local Economic Growth. (2016) *Evidence Review 3: Sports and culture*. London: WWCLEG.

What Works Centre for Local Economic Growth. (2016) *Evidence Review 10: Area-Based Initiatives: Enterprise Zones*. London: WWCLEG.

What Works Centre for Local Economic Growth. (2015) *Evidence Review 5: Estate renewal*. London: WWCLEG.

What Works Centre for Local Economic Growth. (2014) *Policy Briefing: Public Realm*. London: WWCLEG.

7. Appendices

Table 1: Relevant data sources for potential outcome variables

Firm level

Variable	Relevant datasets
Employees	Business Structure Database (BSD)
Labour productivity (sales per employee)	
Total factor productivity	Annual Respondents Database X (ARDx)
Turnover	Business Structure Database (BSD)
Gross Value Added	Annual Business Survey (ABS) / Annual Respondents Database X (ARDx)
Profit	Companies House (CH) / National Business Database (NBD)
Assets	
Formation/Death/Survival	Companies House (CH) / National Business Database (NBD) / Business Structure Database (BSD)
Payroll wages	Annual Business Survey (ABS)
Research and Development	Business Expenditure on Research and Development (BERD)

Individual / Household / Area level

Variable	Relevant datasets
Economic status (e.g. employment)	Annual Population Survey (APS) Quarterly Labour Force Survey (QLFS)
Income, hours and earnings	Quarterly Labour Force Survey (QLFS) Annual Survey of Hours and Earnings (ASHE) Google Street View (GSV)
Outstanding loans	SME and personal outstanding loans from BLGDRC
Deprivation	Index of Multiple Deprivation (IMD)
Subjective well-being	Annual Population Survey (APS)
Crime	Recorded Crime Survey (RCS)
Mobility rates	Annual Survey of Hours and Earnings (ASHE)

Variable	Relevant datasets
Demographic change (university degree, ethnic group, total population)	UK Census of Population (Census) Annual Population Survey (APS) Quarterly Labour Force Survey (QLFS)
Population	UK Census of Population (Census)
Property prices	Price Paid Data from the Land Registry (LR)
Rents / land values / commercial property	Property listings on online platforms such as Zoopla, UK Land Directory, Rightmove, etc.
Various geographic outcomes	GeoData packs from CDRC
Redevelopment / Physical environment	Social media platforms such as Twitter or Flickr.

Table 2: Overview of relevant datasets

Dataset	Source	Access	Cost	Smallest geography	Fine scale analysis?
Annual Business Survey (ABS)	UKDS	Secure dataset	No fee	Postcode	No
Annual Population Survey (APS)	UKDS	Secure dataset	No fee	OA	Maybe
	NOMIS	Public dataset	No fee	LAD	No
Annual Respondents Database X (ARDx)	UKDS	Secure dataset	No fee	Postcode	No
Annual Survey of Hours and Earnings (ASHE)	UKDS	Secure dataset	No fee	OA	Maybe
	NOMIS	Public dataset	No fee	LAD	No
Business Structure Database (BSD)	UKDS	Secure dataset	No fee	Grid coords	Yes
Business Expenditure on Research and Development (BERD)	UKDS	Secure dataset	No fee	Postcode	No
Companies House (CH)	Companies House	Web API* – Public	Limited free access	Postcode	Yes
GeoData Packs	CDRC	Public dataset	No fee	Geodata	Yes
Google Street View (GSV)	Google Maps	Website* – Public	No fee	Grid coords	Yes
Index of Multiple Deprivation (IMD)	Neighbourhood Statistics	Public dataset	No fee	LSOA	Yes
National Business Database (NBD)	BLGDRC	Safeguarded	No fee	Postcode	Yes

Dataset	Source	Access	Cost	Smallest geography	Fine scale analysis?
Outstanding SME loans and outstanding personal loans	BLGDRC	Safeguarded	No fee	Postcode sector	Yes
Price Paid Data	Land Registry	Public dataset	No fee	Postcode	Yes
Property listings online	e.g. Zoopla	Web API* – Public	Limited free access	Grid coords	Yes
Recorded Crime Survey (RCS)	GLA Datastore	Public dataset	No fee	LSOA	Yes
Social Media	e.g. Flickr, Twitter	Web API* – Public	Limited free access		
Quarterly Labour Force Survey (QLFS)	UKDS	Secure dataset	No fee	OA	Maybe
UK Census	UKDS	Public dataset	No fee	OA	Yes

Notes for Appendix Table 2

NOMIS is an ONS-funded online platform for UK small area labour market and socio-economic statistics. Basic access is free although registration is required.

<https://www.nomisweb.co.uk/>

UKDS (UK Data Service) is an ESRC-funded online repository of UK and international social and economic data, with a broader range of datasets than NOMIS and typically at more detailed spatial level. Some UKDS datasets are open and free to download from the site. Other datasets are restricted, and are downloadable by trained researchers after an application process. The most detailed datasets are fully secure, and only accessible by trained researchers through the Secure Lab (SDS) online environment. Restricted and secure access datasets are free for academic use and for public good research, but there are daily use charges for commercial applications. <https://www.ukdataservice.ac.uk/>; <https://www.ukdataservice.ac.uk/use-data/secure-lab/>.

The BLGDRC (Business and Local Government Data Research Centre) and CDRC (Consumer Data Research Centre) are both ESRC-funded online data repositories. They were set up as part of phase 2 of the Big Data Network to make data available to academics, businesses and local government. Safeguarded data require the submission of a Data Access Request form. <http://www.blgdataresearch.org/>; <https://www.cdrc.ac.uk/>*

Whilst websites and web APIs are publicly available there are usually restrictions on access (e.g. 500 calls per minute for Companies House API). Further, collecting data from websites and APIs requires some programming knowledge.