Department for Transport

Transformational Impacts

Extension to West Coast Main Line Case Study

Final 12/08/2022

Contents

Tables

Figures

Appendices

Executive Summary

The Department for Transport (DfT) commissioned CEPA and Arup to conduct an assessment of certain economic impacts related to the upgrade of the West Coast Main Line (WCML) that occurred between 2002 and 2009. This upgrade delivered reductions of up to 25% in effective journey times along the line and by 2016 passenger journeys had nearly doubled from pre-upgrade levels. The economic impacts resulting from the upgrade were assessed at the local authority level, and focused on Gross Value Added (GVA), employment, population, and house prices. In addition the effect on sectoral GVA shares, labour force participation, local authority age profiles and education levels were also examined.

These measures were selected as employment and population increases are indicative of improved connectivity, attracting more firms and people to areas that may have benefited from the upgrade, and leading to higher levels of economic activity. Higher GVA and productivity effects may arise due to greater spillovers from increased economic activity and better firm-employee matching enabled by the increased connectivity. All else equal, these economics benefits would also be expected to be reflected in higher house prices. Finally broader economic opportunities may also be reflected in a younger and more educated workforce, greater labour market participation, and changes in sectoral share towards newer industries.

We assessed potential impacts of the WCML upgrades from 2005 to 2016 by applying a panel difference-in-difference methodology. The method compares outcomes in UK local authorities with WCML stations, to outcomes in similar local authorities without WCML stations, while controlling for changes to other relevant economic and demographic factors related to the local authorities and overall trends in the UK economy and its regions. This is a widely used and well understood econometric method.

In line with initial expectations, local authorities with a WCML station experienced a positive significant difference in the GVA, employment and population following the upgrade. There were also positive significant differences for the share of young adults residing in these local authorities and the share of output in the services sector. Contrary to expectation, we found no significant difference in house prices over the time period analysed and a limited impact on other sectoral and demographic characteristics.

Economic growth is likely to be focused in major cities, and could relate both to the WCML upgrades or to wider economic trends. To further examine the results, we extended the analysis to include a separately estimated effect for major cities along the WCML. This analysis revealed those major cities as a driver of the significant initial results. However, it also revealed that economic trends in these major cities were diverging from other WCML local authorities prior to the upgrade. Therefore, including those cities makes it more difficult to attribute the impacts to the WCML upgrades. Subsequent analysis omitting those major cities however, still indicated significant positive effects for GVA, population, employment and the share of services following the upgrade, though only after completion of the final phase of the upgrade (post 2009).

Finally, a robustness analysis using the ECML local authorities was conducted to further probe the significance and causality of the WCML upgrade. Substituting the ECML local authorities for the WCML local authorities in the above analysis it was found that local authorities along the ECML experienced similar economic outcomes following the WCML upgrade – except for the share of services sector. It is therefore difficult to attribute the positive economic outcomes estimated along the WCML to the WCML upgrade directly. However, this does not mean the WCML upgrade did not produce them, reasons we have not identified causality may be because:

- The upgrade was in part a catch-up on foregone maintenance and improvements. Had the upgrade not happened, the economic outcomes of WCML local authorities may have compared less favourably to similar local authorities, such as those on the ECML. The WCML had high passenger growth numbers compared to the ECML; therefore, the

upgrade may have provided an improved foundation for WCML areas to take advantage of pre-existing trends of population and economic activity shifting to urban areas. In the absence of the upgrades, the WCML local authorities may have been less able to take advantage of these pre-existing trends and fallen behind. That is to say that due to lack of investment, all else being equal, the correct counterfactual may have been for WCML regions to fall behind the ECML instead of maintaining the same trend as indicated in the robustness analysis. Further work should also more rigorously test how appropriate a control group the ECML is. For example, the behaviour of house prices, particularly in major cities, would appear to differ from the experience on the WCML.

- Although the approach taken by this work has been used in academic literature, this available data meant that a relatively simple definitions of treatment area was used: i.e. a local authority with or without a WCML station. Measuring impacts at local authority level may have included areas that were further from a WCML station and omitted the effects of the WCML on local authorities adjacent to WCML stations. This may have downwardly biased the results. The approach also did not account for the level of journey time savings at a specific station. To overcome this, a more rigorous level of exposure to the WCML upgrade could be used, either by looking at improvements in journey times by local authority, and/or by using more granular spatial areas built up from smaller geographic units. This might produce more precise estimates of economic impacts. Other more sophisticated methods such as the Synthetic Control Method that are specifically designed to overcome situations when key difference-in-difference assumptions fail may also be worth exploring.

In conclusion this analysis found some evidence of improved economic outcomes along the WCML following the upgrades completed between 2005 and 2009 but was not able to robustly link those improved economic outcomes to the upgrade. Further work is needed to more rigorously identify the effect of the upgrade within each region.

e jamie.lenney@arup.com

Arup Group Limited

8 Fitzroy Street

London

W1T 4BJ

United Kingdom

1. Introduction

1.1 Project purpose and scope

This report is designed to supplement the West Coast Main Line (WCML) case study from the "Understanding transformational impacts of transport interventions" project that CEPA and Arup jointly undertook for the Department for Transport (DfT).

We agreed with DfT to use a panel fixed effects econometric method to study the impact of the WCML on employment and employment sectoral shares using local authority level data (Annual Population Survey). The analysis was extended to consider local authority GVA, population, house prices and other demographic variables[1] of interest. Due to data availability issues[2], GVA sectoral shares were used instead of employment sectoral shares; they are however closely[3] related.

1.2 Theory and Related Literature

Large inter-regional transportation investments, such as major rail projects, are justified on the grounds of their ability to deliver journey time savings, decongestion, environmental benefits and economic regeneration.

The rationale for rail infrastructure improvements (inputs) is that they lead to reduced journey times, higher services frequencies, and higher capacities across the route (outputs). The increased accessibility that such upgrades afford may then increase the attractiveness of locations on the route, for example as places to live and work, places to visit and for businesses to locate and invest (outcomes). In turn, population and employment opportunities can be expected to increase alongside productivity and a variety of services in a virtuous cycle (impacts). The academic underpinning of this theory is outlined by Redding and Turner.[4]

Broader economic opportunities may also be reflected in a younger and more educated local workforce, higher levels of labour market participation, and sectoral reallocation towards newer industries. For example (in the opposite direction), Gibbons et al (2018)[5] studied the implications of the Beeching Axe[6] and showed that loss of access by rail caused a relative population decline, a decline in the proportion of skilled workers, and decline in the proportion of young people in affected areas.

This study on the WCML fits into the limited, but growing, literature that attempts to assess the evidence of rail schemes to deliver economic growth and regeneration. The What Works Centre for

[1] The share of the population with an NVQ4 qualification or higher, the labour force participation rate, the share of the population in manufacturing and share in services, the share of the population aged 26-39 and aged over 65.

[2] The Business Register and Employment Survey and Annual Business Inquiry are employer-based surveys of locational employment. The BRES replaces the ABI in 2009. Inconsistency surrounding local authority boundary definitions and groupings between the surveys as well as inconsistent industry groupings has prevented us from using local authority level location-based employment data.

[3] GVA is the product of labour and capital with labour compensation representing around two thirds of GVA on average. Compensation represents the product of the wage and hours worked.

[4] Redding, S. J. and Turner, M. A. (2015). Transportation Costs and the Spatial Organization of Economic Activity. Handbook of Regional and Urban Economics, 5:1339–1398.

[5] Gibbons, S., Heblich, S., & Pinchbeck, T. (2018). The spatial impacts of a massive rail disinvestment program: the Beeching Axe.

[6] The Beeching reports published in 1963 recommended the closure of thousands of stations across Britain and resulted in the decommission of thousands of miles of railway lines in the 1960's and 1970's.

Local Economic Growth (2021) conducted a literature review[7] on rail investments where it identified a notable evidence gap on local economic impacts of rail investment, namely employment, demographic composition, rents and property prices. The review of the existing literature indicated some evidence for employment and commercial effects in the immediate surrounding of stations. Studies also tended to find positive effects on property prices, although the size and distance with which they occur varied across studies.

Studies such as Ahlfeldt and Feddersen (2018)[8] find that the GDP of three counties with intermediate stops between Cologne and Frankfurt increased by 8.5% due to the opening of the German high-speed rail network. Coronado and Urena (2018)[9] find that there is evidence that the introduction of Spanish high-speed rail increased spatial inequality by creating winner and loser areas and led to large differences in accessibility. Still, the overall impact of Spanish high-speed rail between Madrid and Seville is likely to have been positive. Employing similar methods to this study, Carbo et al (2019) found that provinces with stops on the Madrid-Barcelona HSR line experienced a 2.4% increase in economic output, 3.3% increase in the number of firms, and 1.1% increase in labour productivity, in real terms.

2. WCML Upgrade

This study focuses on the 2002-2009 WCML upgrade. The WCML is one of the busiest railway corridors in the UK, running from Euston station in London, touching several UK regions and connecting the major cities of Birmingham, Manchester, Liverpool, Glasgow, and Edinburgh (Figure 1).

The network is a mixed-traffic railway, providing intercity, regional, commuter and freight rail services and has existed in some form for approaching 200 years. A brief history from 1830 (when the line construction began) to 2001 (when Railtrack went into administration) is provided in Box 1.

Box 1 – Brief history of the WCML

Construction of the line began in 1830 and was completed in 1837. The final sections of what is now known as WCML were opened in 1869. In January 1923, the route came under the control of the London, Midland and Scottish Railway, which competed fiercely with the East Coast Main Line (ECML) to minimise journey times during the 'Race to the North'. In 1948, the service was nationalised and the line was officially referred to as the WCML. Electrification proceeded between 1959 and 1974.[10] By the 1990s, the route required substantial investment due to aging infrastructure and service unreliability. Railtrack, the privatised predecessor of Network Rail, commenced an upgrade of the line in 1996 and then in 1998 made a private agreement with Virgin Trains for a more ambitious upgrade.[11] These plans did not proceed well and were interrupted by the Hatfield crash. Railtrack subsequently went into administration in 2001.

[7] What Works Centre for Local Economic Growth (2021). Rapid evidence review: Rail investment: Available at: https://whatworksgrowth.org/public/files/Policy_Reviews/Rapid_evidence_review_Rail_investment_What_Works_Growth.pdf (Accessed: 26 June 2022)

[8] Ahlfeldt, G. M. and Feddersen, A. (2018). From periphery to core: measuring agglomeration effects using high-speed rail. Journal of Economic Geography, 18(2):355–390.

[9] Coronado, J. M. and Urena, J. M, (2018). Long term implications of HSR on small cities: Ciudad Real and Puertollano revisited 25 years after the arrival of HSR. 360.high speed magazine, 6:207-223.

[10] British Transport Commission (1955). Modernisation and Re-Equipment of British Railways. Available at: https://www.railwaysarchive.co.uk/docsummary.php?docID=23 (Accessed: 12 April 2022).

[11] House of Commons Library (2010). Railways: West Coast Main Line, Louise Butcher. Available at: https://commonslibrary.parliament.uk/research-briefings/sn00364/ (Accessed: 12 April 2022).

Figure 1: The West Coast Main Line (WCML)

Source: ONS and author calculations.
Note: Dark red highlighted local authorities are local authorities that contain a WCML station (the treatment group in our study). Local authorities in London are excluded from the analysis (see section 3.1).

Network Rail took over the WCML project from its predecessor Railtrack and concluded that the ambitious plans for the line were undeliverable. In 2002, the Strategic Rail Authority intervened and set the following objectives for a revised upgrade:

1. Address the major backlog of maintenance and renewals on the route, ensuring value for money;

2. Provide an improved level of performance, safety and reliability, which will in turn help the railway regain lost market share and increase the role it can play in the national and regional economies;

3. Provide capacity for anticipated growth in passenger and freight business over the next 20-30 years, with substantially faster and more competitive journey times;

4. Establish sustainable and cost-effective maintenance regimes; and

5. Achieve these objectives on a 'working railway' while allowing for the continuation of key freight and passenger traffics during the rebuilding and enhancement work.

The primary objective of the scheme was to address the backlog of maintenance and renewal works, whilst improving journey times and reliability.

The WCML upgrade spanned the entire route between London and Glasgow, with connections to Birmingham, Liverpool, Manchester, Edinburgh and Glasgow and was completed in three phases:

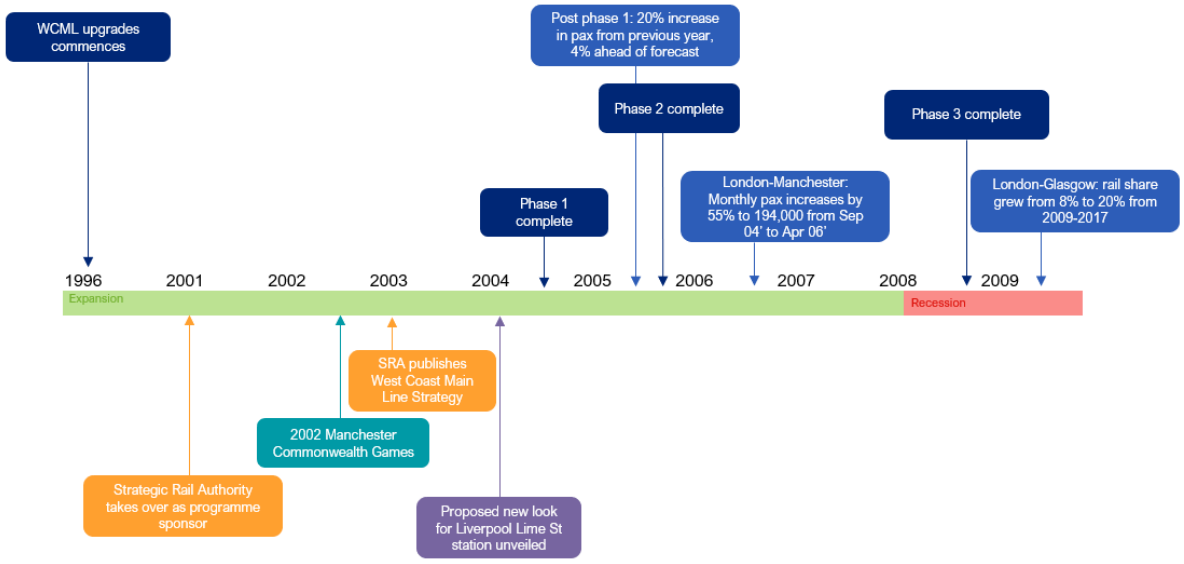
1. Phase 1 September 2004: Track upgrades to 125mph between London, Manchester, Birmingham, and Crewe to produce a more frequent timetable;

2. Phase 2 December 2005: Track upgrades to 125mph between Preston and Glasgow; and

3. Phase 3 December 2008: Renewals and enhancements to increase long-distance train capacity by 80% percent and to provide 70% more freight paths compared to the period pre-2004.

Figure 2 below presents the timeline of the WCML upgrade. The phasing of work resulted in a staged delivery of service changes and benefits to rail users. Figure 3 illustrates WCML phasing over time by plotting the change in generalised journey time (GJT) for the Glasgow to London and Manchester to London route.[12] For context the GJT between Newcastle and London on the ECML is also plotted, showing that it did not improve noticeably over the period. GJT[13] combines multiple features that make a journey more or less attractive, such as improvements in on-train journey time, as well as improvements in timetabling and train frequency.

Figure 2: Timeline of the West Coast Main Line Upgrade



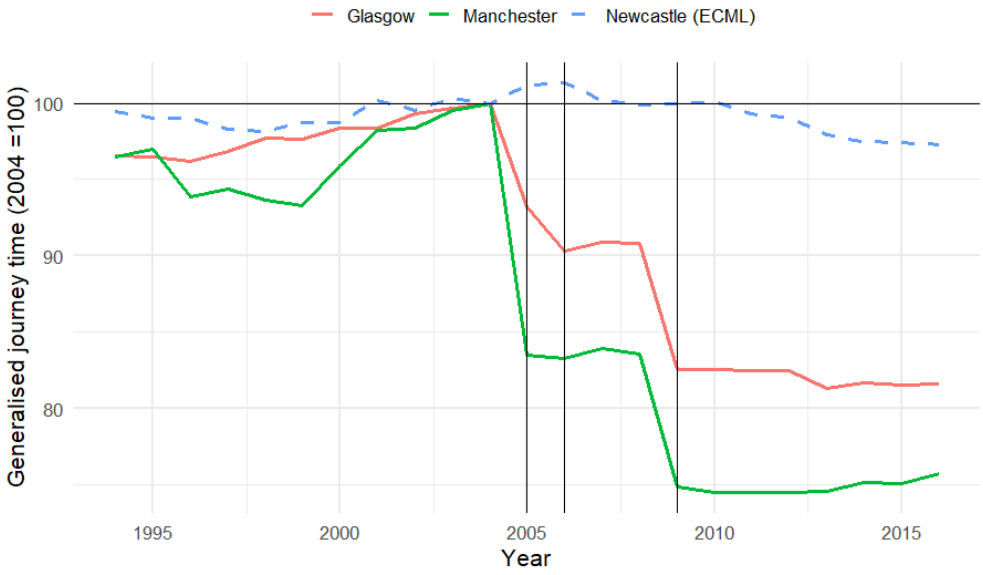
Source: Transformative Impacts of Transport Investments: Case Study Report. CEPA and Arup (2022).

Figure 3 shows a moderate increase in GJT leading up to 2004, implying a deterioration in journey experience. Based on interviews with two DfT staff working on the WCML upgrade at the time, this deterioration reflected a backlog of maintenance, speed restrictions on the rail network following the Hatfield rail crash and construction related to the upgrade work itself. Upon completion in 2005, WCML Phase 1 delivered around a 15 percent reduction in GJT for the Manchester-London route.

[12] For more information on journey time reductions across the line see appendix Table A.1.1.
[13] Department for Transport (2021). TAG Unit A.1.3: User and Provider Impacts. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1007443/tag-unit-A.1.3.pdf (Accessed: 12 April 2022).

The following year, Phase 2 delivered a 10 percent reduction in GJT along the whole route (Glasgow-London). In 2009, Phase 3 brought a final improvement in GJT resulting in an overall reduction of 25 percent between Manchester and London, and around 20 percent between Glasgow and London. This represented a significant line-wide improvement and led to robust growth in passenger volumes in excess of the business case forecast for that year.[14]

Figure 3: Generalised Journey Time Changes on the WCML

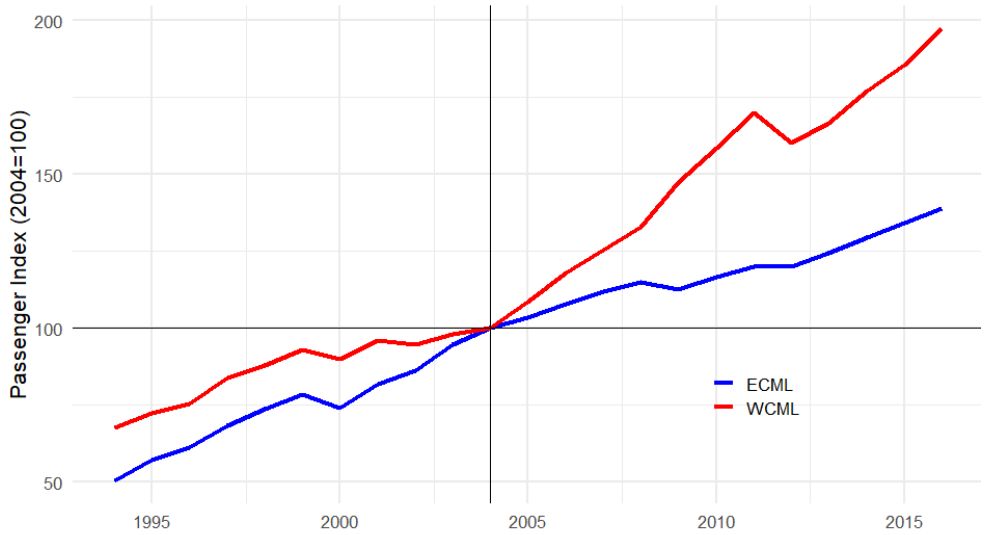


Source: DfT (RUDD database)
Note: Chart shows the generalised journey time for the Glasgow-London route and Manchester to London route indexed to their 2004 value. The dashed blue line shows the generalised journey time for the ECML Newcastle-London route. Solid black lines indicate when phases were completed.

Figure 4 shows the number of recorded journeys from WCML stations and ECML, with growth on the WCML accelerating following the upgrade relative to growth on the ECML. Passenger volumes grew by around 20% following the completion of Phase 2 and more than doubled by 2016. Some of this additional growth was attributed[15] to mode shift from air to rail following the upgrade; flights to Manchester from other cities along the WCML peaked in 2004 and the market share trips between London and Manchester of rail rose to 85% in 2017 from around 60% in 2009. These numbers suggest[16] that, absent of this mode shift, rail travel along the WCML would have grown by around 50% instead of the 100% recorded in Figure 4. Therefore, the upgrade is likely to have delivered a substantial economic benefit in terms of emissions reductions (23,000 Tonnes between 2009 and 2017[15]).

[14] National Audit Office (2006). The Modernisation of the West Coast Mainline. Available at: https://www.nao.org.uk/report/the-modernisation-of-the-west-coast-main-line/ (Accessed: 12 April 2022).
[15] Campaign for Better Transport (2019). Transformation of the West Coast Mainline. Available at: https://bettertransport.org.uk/sites/default/files/research-files/Transformation-of-the-West-Coast-Mainline.pdf (Accessed: 26 June 2022).
[16] Holding total trips constant in Figure 4 but assuming market shares remained unchanged between rail and aviation from 2004.

Figure 4: Annual Passenger Traffic on the WCML



Source: DfT (RUDD database)
Note: Chart shows total passenger journeys originating from stations on the WCML and ECML indexed to their value in 2004. Passenger Journeys with distances of less than 30 miles and journeys within London and the SE have been excluded.

3. Data and Methodology

This analysis uses publicly available representative data at the local authority level to estimate the economic effect of the WCML upgrades on selected variables of interest. Effects are estimated using a difference in difference regression that compares the average level of economic variables of interest (e.g. employment) in local authorities affected by the upgrade to the average levels of similar local authorities that were not affected – the difference between the two being the estimated effect. Similar local authorities were defined as local authorities within the same NUTS1 region[17], with similar demographic characteristics that do not have a WCML station. We estimated separately the effect for major cities relative to other local authorities. Finally, to test the robustness of our results, we conducted a placebo test whereby we falsely assumed that the ECML (instead of the WCML) received an upgrade over the same time period and estimate effects for the ECML.

3.1 Data sources and summary statistics

Data is used from the following sources at the local authority district level for the period 1999-2016:

1. ONS Annual Population Survey (APS) [18]
2. ONS regional GVA estimates[19]
3. House Price Statistics for Small Areas (HPSSAs)[20]
4. DfT Rail Usage and Drivers Dataset (RUDD)

The **APS** provides representative economic and demographic information at the local level on an annual basis for the UK[21]. It provides data on population, employment and unemployment, as well as housing, ethnicity, religion, health and education. We have supplemented this dataset with local authority house price data from the HPSSAs, annual sectoral GVA data by local authority from the ONS's national accounts, as well as journey time and passenger statistics from the RUDD dataset, which we have aggregated to the local authority level.

The **GVA** data provides estimates for gross value added (GVA) by local authority and year, in total and by industry. This data is constructed from the components of balanced NUTS3 level GVA data and apportioned to the local authority level using different ONS surveys. Specifically, the Business Register and Employment Survey (BRES) is used to apportion employment income using locational employment data, the Annual Business Survey is used to apportion capital income and local housing values are used to apportion rental income.

[17] North East, North West, Yorkshire and the Humber, East Midlands, West Midlands, East of England, Greater London, South East, South West, Wales, Scotland. See: https://www.ons.gov.uk/methodology/geography/ukgeographies/eurostat (Accessed: 06/08/2022)

[18] Office of National Statistics. Annual population survey (APS) QMI. Available at: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/methodologies/annualpopulationsurveyapsqmi (Accessed: 12 April 2022).

[19] Office of National Statistics. Regional gross value added (balanced) by local authority in the UK. Available at: https://www.ons.gov.uk/economy/grossvalueaddedgva/datasets/regionalgrossvalueaddedbalancedbylocalauthorityintheuk Accessed: 12 April 2022).

[20] Office of National Statistics. House price statistics for small areas QMI. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/housing/methodologies/housepricestatisticsforsmallareasqmi#quality-characteristics-of-the-data (Accessed: 12 April 2022).

[21] We've excluded Northern Ireland from our analysis as there is no land connection between Northern Ireland and the other regions and Northern Ireland is subject to the Eurozone's business cycle which differed significantly over the period in question. We also exclude the Scottish islands, the Isle of Scilly and the City of London.

The **HPSSA** is built from administrative land registry data on housing transactions in England and Wales and therefore excludes Scottish transactions. The dataset we use reports the median house price by local authority.

The **RUDD** data is a database maintained by DfT to assist with forecasting rail passenger demand. It contains data for station pairs (origin and destination) on the UK rail network. Information includes passenger numbers; generalised journey time that takes account of frequency of service as well as actual journey times; and, other variables of interest such as journey distance.

In our analysis we excluded local authorities from Northern Ireland, the Scottish Islands and the Isles of Scilly, as they were thought to be sufficiently disconnected from the rest of the UK's economic cycle (for example, the Northern Ireland local authorities are subject to spill overs from the Eurozone business cycle). We also excluded affected local authorities in London due to concerns over significant economic spill overs between them, which make determining upgrade impacts difficult.

Table 1 averages selected variables of interest by local authority for the 'pre-treatment period' before the upgrade (1999-2004) and the 'post treatment period' (2010-2016). The table provides information for local authorities associated with stations on the WCML, ECML and non-WCML local authorities (rest of UK including the ECML). A local authority is associated with a main line if a main line station exists within its geography (see Appendix Table A.1.1 for a list of ECML and WCML stations). We included the ECML in Table 1, as it is used as a robustness test for our results.

Table 1: Averages of selected Local Authority Characteristics

	WCML		ECML		Rest of UK*	
	1999-2004	**2010-2016**	**1999-2004**	**2010-2016**	**1999-2004**	**2010-2016**
Population	175,187	192,413	171,277	188,212	106,225	116,210
Population Growth (%)	0.7	0.7	0.6	0.8	0.7	0.7
GVA per head	16,538	22,784	16,803	22,489	15,580	21,080
GVA per head Growth (%)	4.4	2.8	4.7	2.2	4.1	2.4
NVQ4 Share (%)	23.3	32.5	23.8	33.2	23.6	33.0
Economic Activity Rate (%)	63.3	62.8	63.8	63.1	62.9	62.5
Unemployment Rate (%)	5.1	6.5	4.5	6.6	4.4	6.0
Manufacturing share GVA (%)	16.7	12.8	15.0	12.6	16.6	13.6
Services share GVA (%)	17.2	19.7	18.3	19.3	15.6	17.4
Public Sector share GVA (%)	16.9	19.2	18.1	21.0	17.5	20.0
House Price	94,230	168,703	95,839	174,635	106,802	191,166
House Price Growth (%)	14.6	3.2	14.9	3.5	14.7	3.7
Population aged 26-39 (%)	20.5	17.8	20.3	17.5	19.7	16.5
Population 65+ (%)	14.2	16.4	14.7	16.7	15.4	18.0
No. Local authorities	43		20		298	

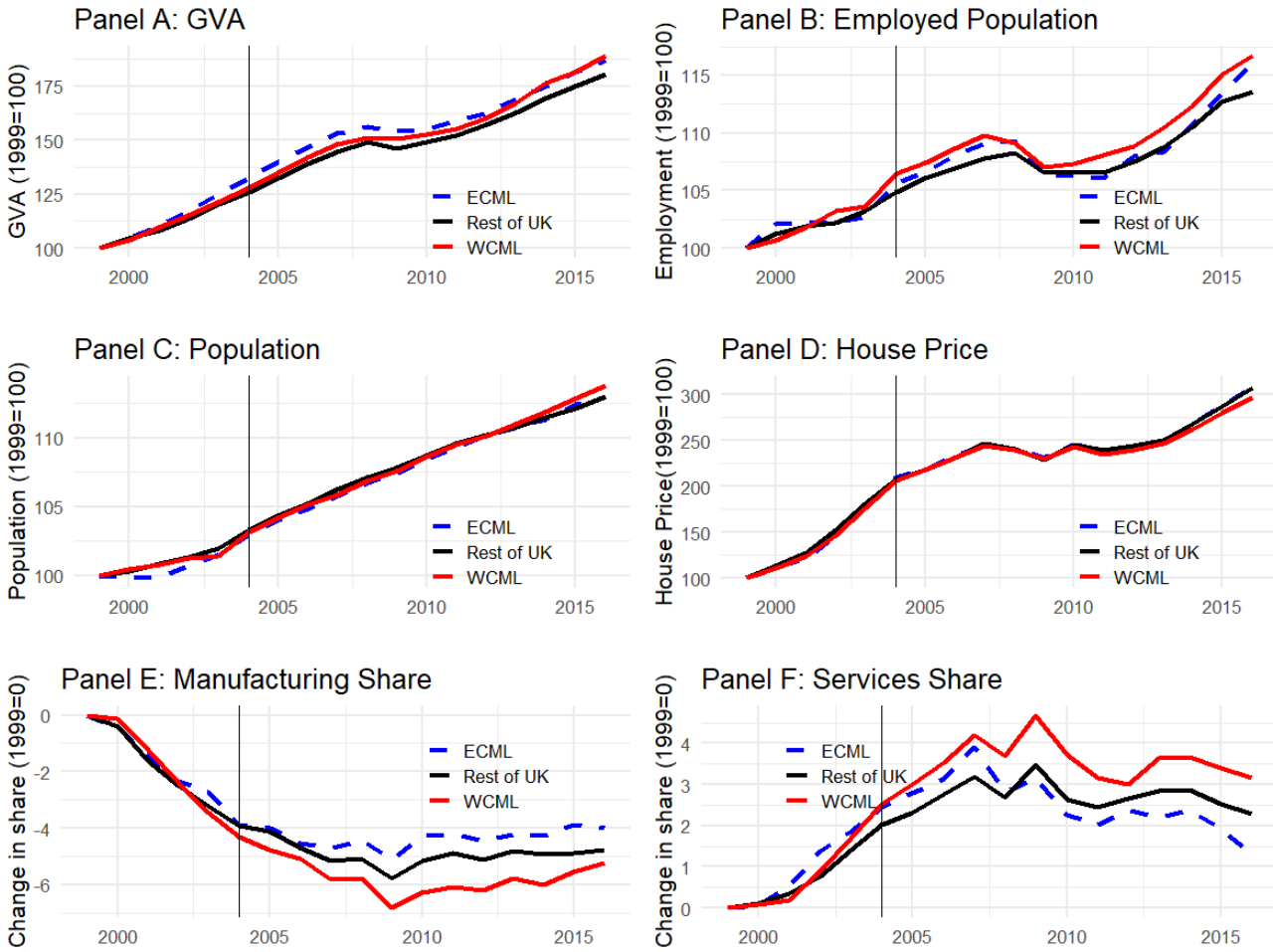
Source: ONS, UK Land Registry
Note: Simple averages by local authority and year shown. London, Northern Ireland, Scottish Islands, and the Isles of Scilly are excluded. Simple averages are consistent with fact regressions are not weighted by population.
*Rest of UK includes ECML.

We see that both local authorities with stations on the WCML and ECML had on average larger populations than the average local authority, which is not surprising given they include many large urban centres. Consistent with that, the WCML and ECML local authorities have a slightly younger adult population as reflected by the age share over 65. They also have slightly higher GVA per head, which will in part reflect productive industries located in city centres. It may also reflect local

authorities along these lines having a higher employment share made up of commuters, who increase local GVA but not local population. On average, the WCML local manufacturing share fell more than the rest of the country between the treatment periods. This is also likely reflected in the dynamics of the WCML's services sector, which grew faster than the rest of the country. The WCML local authorities also have a slightly lower public sector share of GVA and slightly lower house prices.

On average, WCML local authorities do not appear to be significantly different from ECML local authorities or from the rest of the UK. This is shown in Figure 5, which plots the dynamics of some of these variables over our sample period. We see this particularly in the pre-upgrade period, when the dynamics of the variables of interest are observationally similar between the WCML and the rest of the UK except for the share of services GVA.

Figure 5: Economic Dynamics Pre and Post Upgrade



Source: ONS & Arup calculations
Note: Chart plots averages of local authorities for the WCML authorities, ECML authorities and non-WCML local authorities for six variables. Panels A-D are indexed to the 1999 values. Panels E and F shows the change in manufacturing share of GVA relative to 1999 in percentage points. WCML/ECML local authorities are local authorities containing a WCML/ECML station. See Appendix Table A.1.3 for a list of local authority locations.

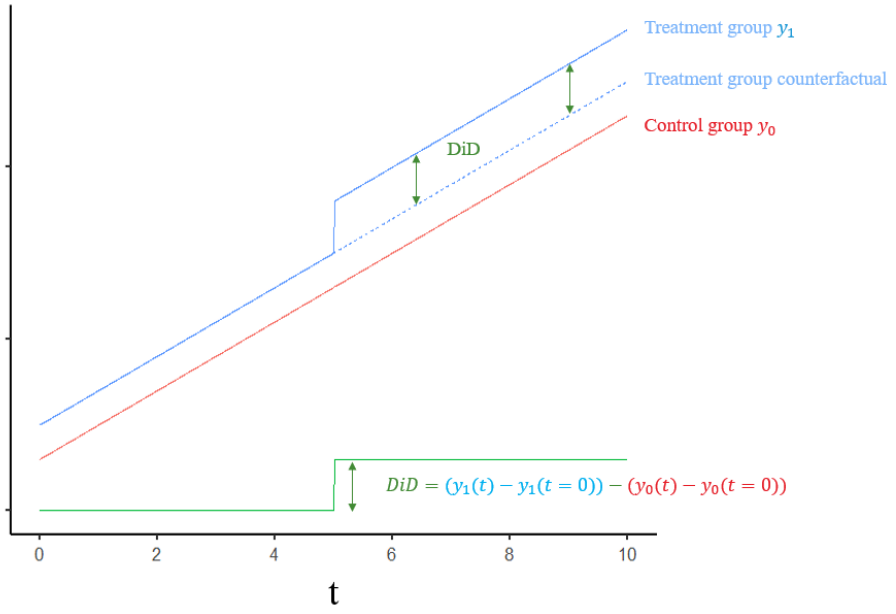
3.2 Difference in Difference model

3.2.1 General model

Our main approach relied on the difference in difference (DiD) specification[22] which is one of the most prevalent ways to estimate the effects of a policy intervention. In DiD analysis, we refer to a "treatment group" which experiences an intervention, in this case the railway upgrade. It is compared against a "control group", which does not experience an intervention. The size of the difference is known as the "treatment effect".

The DiD method compares the changes in outcome values before and after an intervention for both treatment and control groups. The before-and-after change in the control group serves as a counterfactual for the treatment group. A DiD analysis is illustrated in Figure 6 which shows the trajectories for two groups for a variable of interest over time. Each group $\{0,1\}$ has the same time trend (common time fixed effect), however there are differences between the groups in terms of their overall level (individual fixed effects). At time period 5 there is an intervention that affects the treatment group, causing the level of the variable of interest y to shift upwards in the treatment group. Absent that intervention, we assume the path of the treated group would have followed the path of the blue dashed line, i.e. the treatment group would have continued to share a common trend with the control group. This is called the common trend assumption. Given this assumption, the treatment effect (shown in green) can be estimated by comparing the changes over time of the two groups.

Figure 6: Illustration of the Difference in Differences with Fixed and Time Effects



In practice, we estimated a fixed effect regression which controls for other factors that could confound the common trend assumption and takes the following form in words to derive a treatment effect:

[22] Hansen, B. E. (2022). Econometric. Available at: https://www.ssc.wisc.edu/~bhansen/econometrics/Econometrics.pdf (Accessed: 12 April 2022).

$$y_{i,t} = individual\ effect_i + time\ effect_t + \delta\ treatment\ status_{i,t} + \beta\ other\ variables_{i,t} + other\ unobserved\ factors_{i,t} \quad (1)$$

Where *i* refers to an individual local authority and t refers to a specific year. Looking at the variables, *y* is a variable of interest like GVA or Employment, the individual effect is the difference in y attributed to specific unobserved local authority characteristics, the time effect controls for general changes in the variable of interest over time, treatment status is a binary indicator and refers to whether the local authority has received the rail upgrade or not, and other variables control for other important observed local authority level characteristics[23]. These may include, for example, sectoral shares that may change over time and may cause or correlate with other factors that cause the local authority to deviate from the common trend of other local authorities.

In this specification we estimated the individual effects (*i*), the time effects (*t*), the treatment effect (δ) and coefficients for the effects of other variables (β). In our case, the variable of interest $y_{i,t}$ will either be the log of an economic quantity (e.g. GVA) or a sectoral share like the GVA share of manufacturing. The log functional form is chosen when we want to estimate the treatment effect as percentage effect on the level of the variable.

Our estimate of δ empirically answers the question of whether on average local authorities exposed to the WCML upgrade have a statistically significant difference in their outcome for variable y compared to similar local authorities within the region. Unlike the simple instant treatment effect in Figure 6, it is likely that in our case impacts will build over time and also have some local authorities being treated at later dates to others. This means that our estimate is an average[24] of treatment effects over time between treated and untreated groups.

3.2.2 Key Assumptions

Three key assumptions are needed for DiD to robustly estimate δ and be confident of treatment causality. They are:

1. The common trend assumption. That absent the intervention, the two groups would have followed similar trajectories for the variable of interest;

2. that there is a clear cut-off separating the treatment group from the control group; and

3. the treatment group effect will not spill over to the non-treated group.

As we can never truly verify assumption (1) — even in a laboratory setting — we test the assumption holds before treatment and then assume it would continue to hold afterwards. Before estimating equation (1), we conduct a common trend test[25] that empirically tests the validity of the common trend assumption between the treatment and control group in the pre-treatment period. The test estimates trends for the treatment and control groups prior to treatment and asks whether they are different in a statistically significant sense. The null hypothesis is they are not, but if the test rejects the null hypothesis, we have almost certainly violated key assumption (1) as the

[23] Our controls for the major economic variables of interest largely follow Carbo et al (2019) and include the average pre-treatment population (size effect), the share of manufacturing in GVA, share of services GVA, share of public sector GVA, share of those with a degree or higher (NVQ4), the economic inactivity rate and female economic activity rate.
[24] Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. Journal of Econometrics, 225(2), 254-277.
[25] More detail on this test is in Appendix A2 and results of the tests in Appendix A4.

estimated coefficient δ will likely be partially reflecting a continuation of that difference in pre-existing trends.

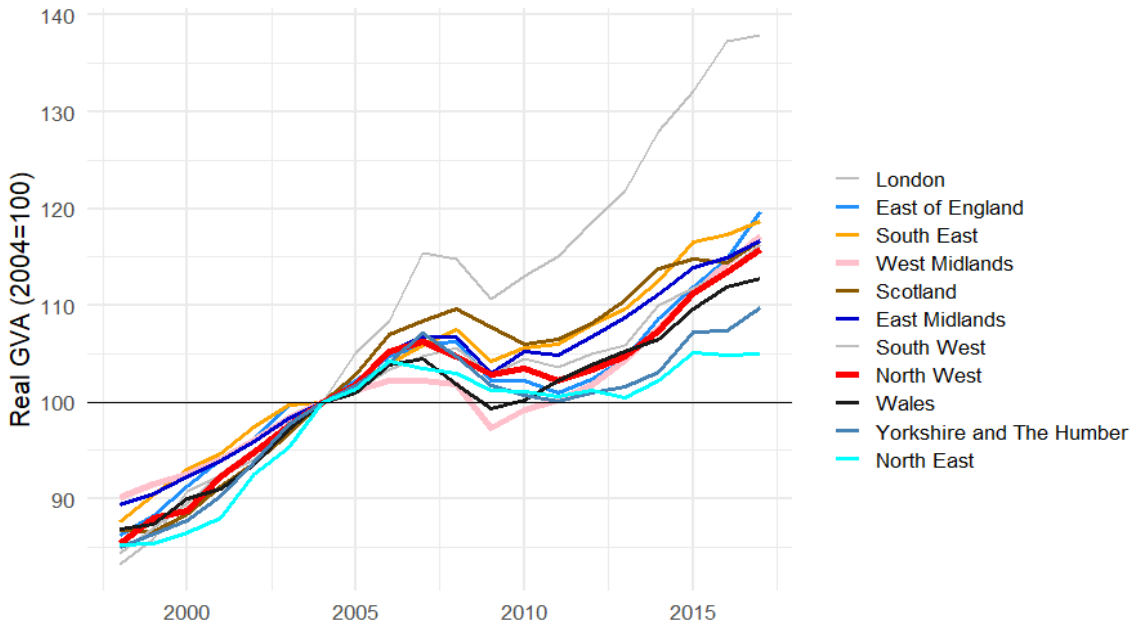
Assumption (2) and assumption (3) are in our case difficult to disentangle and unlikely to hold fully[26] in this analysis as is often the case in the wider academic literature outside of random control experiments. Firstly, there is the issue of defining the treatment. For example, our definition of treatment is that the local authority contains a WCML station serving high speed rail. This means that instances where WCML stations sit near local authority boundaries create an arbitrary cut off in terms of treatment, where in reality both local authorities might be similarly treated. Related to this, it is likely that some local authorities bordering local authorities with WCML stations benefit from positive economic spillovers. This is more pronounced in London, which is the reason why London local authorities were excluded from the analysis. The implications for violating assumptions (2) and (3) are to potentially downward bias our estimates for the absolute treatment effect $|\delta|$, as positive or negative spillovers from one local authority to another will reduce the differential in their economic outcomes and therefore the estimate of δ is conservative.

To estimate equation (1), we included data from all UK local authorities to control for regional time trends. This means our estimate δ is effectively a weighted average of regional treatment effects for regions that have both treated and not treated local authorities[27]. Controlling for regional time trends, particularly around the financial crisis, is important as there exists considerable variation in regional trends and responses to the business cycle. This is shown in Figure 7, which plots GVA by region over our sample period. Over the entire period, regions experienced different growth rates and diverged. There was also considerable variation in the response to the financial crisis, for example, the North East (cyan) suffered a more prolonged slump relative to the rest of the country. A failure to control for specific regional trends, which may not be adequately controlled for by our other control (X) variables, risks conflating any results between their correlation with the WCML and the broader regions it travels through. Variability between regions over time also makes a direct comparison with a natural control group like the ECML difficult, as the two lines have a limited overlap in terms of regions. This makes estimating equation (1) with regional time effects on a sample of only WCML and ECML local authorities invalid from an econometric standpoint. A comparison with the ECML is achieved through a placebo test that estimates a δ for the ECML local authorities assuming the ECML local authorities (instead of the WCML) received an upgrade over the WCML upgrade timeline.

[26] The failure of these assumptions to hold fully is often inevitable outside of scientifically designed random control trials but is less detrimental to our analysis than the failure of assumption (1). Violating assumption (1) results in a complete failure of identification, whereas violation of assumptions (2) and (3) will tend to introduce greater statistical uncertainty into our estimates but does not invalidate statistically significant estimates.

[27] Therefore, this means that local authorities in the South West or North East do not contribute to the estimate of δ as those regions include no treated local authorities.

Figure 7: Regional GVA Trends



Source: ONS
Note: Figure plots regional GVA series deflated by the GVA deflator and indexed to 2004.

3.2.3 Alternative specifications of the model

In addition to an overall average treatment effect, we also estimate treatment effects over time or by different local authority characteristics. This is done by simply augmenting equation (1) to break out the δ coefficient across different indicators. Equation (2) shows this broken out over time, where each time period receives a different estimate. The treatment effect over time is of interest for two reasons. First, because the upgrade was completed in phases. Second, because it may take several[28] years for real economic impacts to be realised.

$$y_{i,t} = individual\ effect_i + time\ effect_t + \sum \delta_t \mathbf{1}_t\ treatment\ status_{i,t} + \beta\ other\ variables_{i,t} + other\ unobserved\ factors_{i,t} \quad (2)$$

We are also interested in whether the upgrade affected major cities differently to other local authorities between major[29] cities. Equation (3) shows the regressions equation relating to this specification, where a separate δ_M estimates any additional effect the upgrade may have had on major (M) cities. This additional effect $\delta_M$sits on top of the treatment effect applicable to all local authorities δ_{LA}, such that the estimated effect for major cities is equal to $(\delta_{LA} + \delta_M)$ and for other local authorities the effect is simply δ_{LA}. This is of interest for several reasons. Firstly, large cities may be better placed to realise the benefits of greater connectivity and become even more attractive to firms and workers. This could be reflected in better economic outcomes for major cities and even potential negative outcomes for other local authorities through displaced economic activity. On the other hand, greater connectivity at intermediate stops may make these places more attractive to live and boost the local economy through a higher commuting population. It is also argued in the literature that, if significant effects are estimated for intermediate stops, these are more credible from a causality standpoint, because rail projects tend to focus on connecting major

[28] For example, construction around stations may only be delivered in the decade following a rail/station upgrade.
[29] On the WCML these cities are Birmingham, Edinburgh, Glasgow, Liverpool and Manchester. On the ECML they are Edinburgh, Leeds and Newcastle.

cities as part of a broader economic strategy. This makes assigning causality to rail more difficult for these major cities that may be benefiting from other investments or underlying trends, whereas the intermediate stops are often argued to be receiving a more plausibly unexpected economic benefit that can be more credibility linked to rail investment.

$$y_{i,t} = individual\ effect_i + time\ effect_t + \delta_{LA}\ treatment\ status_{i,t} + \delta_M\ treatment\ status_{i,,t}\mathbf{1}_{M,i} + \beta\ other\ variables_{i,t} + other\ unobserved\ factors_{i,t} \quad (3)$$

Finally, when testing for significance of the estimated coefficients we adopted a state of the art approach to the estimation of standard errors ("clustering") which recognises the fact that regression residuals ($other\ unobserved\ factors_{i,t}$) are likely to be correlated within individual local authorities over time rather than assuming that residuals are uncorrelated as in a conventional approach to standard error estimation ("robust standard errors"). This is consistent with the assumption that the effects of the upgrade will be heterogenous across local authorities – with each local authority being its own "cluster" and that the reported coefficient, therefore, represents an average of those heterogeneous effects across local authorities. This tends to lead to wider[30], but more accurate and credible confidence intervals when compared to the more common (and in this case incorrect) assumption of arbitrary correlation[31]. When and how to cluster has been subject to considerable debate - for a more detailed explanation, see Abadie et al (2017) [32] – but we regard our approach to be appropriate, prudent and an improvement on comparable literature.

[30] Ozler, B. (2012). Beware of studies with a small number of clusters. Available at: https://blogs.worldbank.org/impactevaluations/beware-of-studies-with-a-small-number-of-clusters (Accessed: 12 April 2022).

[31] This is due to the fact that the regressor of interest δ (e.g., treatment indicator) is usually highly correlated within a cluster (i.e., the treatment variable "WCML" is likely to be either on or off per local authority and it is therefore highly serially correlated). The combination of serially correlated regression residuals and a serially correlated treatment variable makes our results less certain.

[32] Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. (2017). When should you adjust standard errors for clustering? (No. w24003). National Bureau of Economic Research.

4. Results

4.1 WCML

4.1.1 Average Treatment Effect

Table 2 displays the results from our main regression (equation 1) for our variables of interest. We find positive significant[33] estimates for the average difference of the level of WCML local authorities GVA (+2.1%), employed population (+1.6%), population (+1.3%), the share of services (+0.7pp) and the share of young adults aged 26-39 (+0.4pp). We do not find a positive difference for house prices over this period, which is surprising given the other results and further explored in section 4.1.2.

Table 2: Economic Impact (DiD) – Entire Sample

Dependent Variable	Effect (δ)	Parallel Trend (Pass/Fail)
Major economic variables		
GVA (log)	0.021**	Pass
Employed population (log)	0.016**	Pass
Population (log)	0.013*	Pass
Median House Price (log)	-0.016	Pass
Demographic/Sectoral variables		
Manufacturing share	0.000	Fail
Services share	0.748**	Pass
Activity rate	0.101	Pass
NVQ4	-0.006	Pass
Age 26-39 share	0.423*	Pass
Age 65+ share	-0.307	Fail

Note: Table reports results from panel regression with fixed effects on local authority data for the period 1999-2016. The regression includes local authority fixed effects and controls for time trends at the region level. The regression also includes time varying local authority variables. The major economic variables controls include the share of the population with a degree level equivalent qualification (NVQ4), public sector GVA shares, manufacturing GVA shares, service sector GVA shares, the female economic activity rate, the overall economic activity rate and the pre-treatment local authority population. The demographic and sectoral variable controls include the pre-treatment local authority population the log of public sector GVA and the female activity rate. Standard errors are clustered at the local authority level and are also used for the parallel trend test. Region is at the NUTS1 level. Rejection of the parallel trend test implies a failure of the common trend assumption. See Appendix A4, Tables A.4.1, A.4.2, A.4.3 and A.4.4 for more detailed regression results. * = p < 0.1, ** = p< 0.05, *** = p < 0.01

Employment and population increases are indicative of improved connectivity, attracting more firms and people to these areas and leading to higher levels of economic activity. The fact that the GVA point estimates is higher than the employment estimate is indicative of a productivity effect, possibly due to greater spillovers from higher levels of economic activity and the potential for better firm-employee matching enabled by the increased connectivity. The significant effect on the number of young adults aged 26-39 and the share of output in the services sector may reflect the

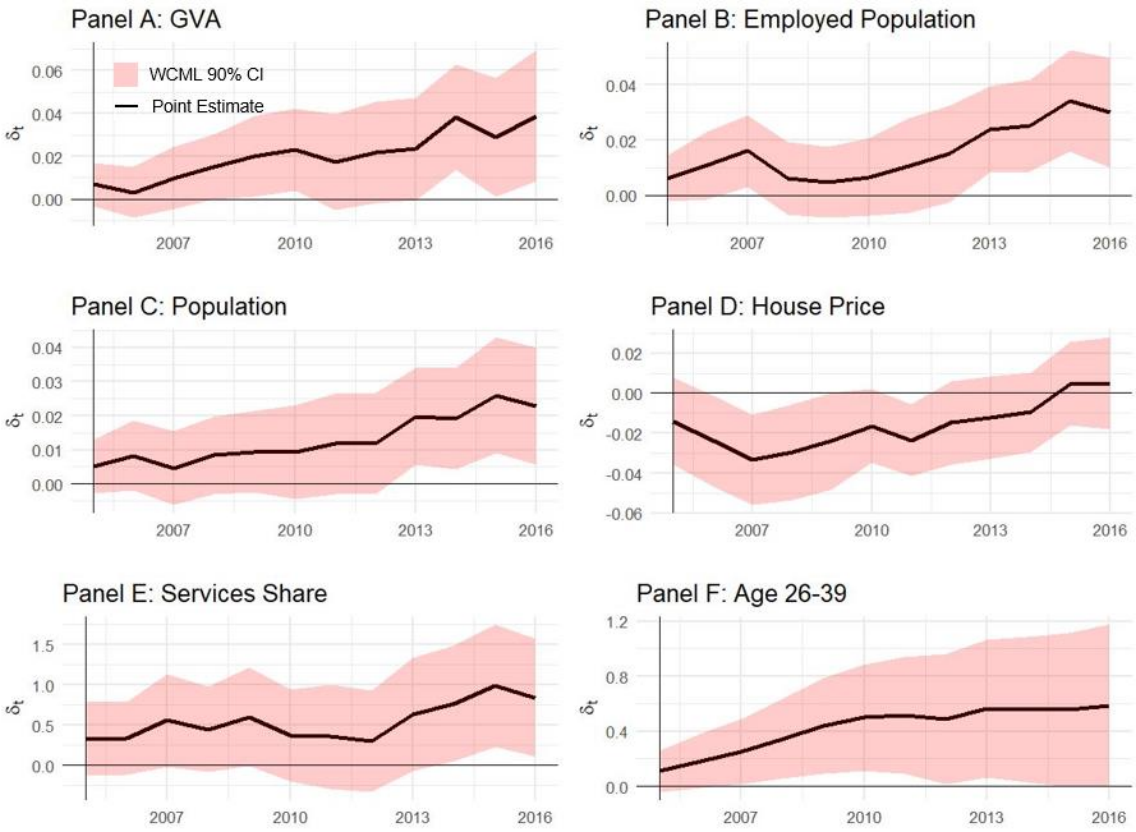
[33] From here on significant means the estimated 90% confidence interval for the variable does not include zero (* = p < 0.1)—a common reporting level. Often a higher 95% confidence interval is selected however given the potential downward biases in our estimates discussed 3.2 from positive spillovers between local authorities, alongside downward bias that may result from measurement error of economic data at the local authority level; that threshold was considered impractically high as the focus of this analysis.

broader economic opportunities afforded to younger adults by enhanced connectivity and an acceleration of the modernisation of the local economies around the WCML through increased connectivity.

4.1.2 Average Treatment Effect Over Time

The results in Table 2 represent an average treatment effect over local authorities and over time. Figure 8 plots the estimate for the effects reported in Table 2 by year, by augmenting the regressions with an interaction between treatment and year (equation (2)). For GVA, employment and population, the point estimates are positive by year, rise over time and estimated to be significantly above zero from around 2013. The house price point estimates are negative at first, significantly below zero in 2007, but this negative estimate fades over time and is indistinguishable from zero by the end of the sample period. The initial negative effect on house prices may reflect the house price slowdown leading up to the financial crisis being greater in urban transit areas, which may have been more likely to experience higher and unsustainable growth and/or an over-supply of new housing during the housing bubble. Finally, the overall gradual increase in the point estimates and significance may reflect a combination of growth effects related to the upgrade and the fact that a significant share of the upgrade was not completed until 2009.

Figure 8: Treatment Effect δ_t for WCML by Year



Note: Figure shows the estimates and 90 percent confidence intervals (red shaded area) for the estimates of δ_t based on the regressions in Table 2 adapted to include time varying treatment effects (equation 2). Standard errors are clustered at the local authority level. Panels A-D are a log difference and Panels E and F are a difference in percentage point share.

4.1.3 Additional Average Treatment Effect on Major Cities

Table 3: Economic Impact (DiD) – Major Cities

Dependent Variable	Effect (δ_{LA})	Parallel Trend (Pass/Fail)	Effect - Major Cities (δ_M)	Parallel Trend - Major Cities (Pass/Fail)	Effect - Excl. Major cities (δ)	Parallel Trend - Excl. Major Cities (Pass/Fail)
GVA (log)	0.015	Pass	0.44*	Fail	0.016	Pass
Employment (log)	0.009	Pass	0.056	Fail	0.01	Pass
Population (log)	0.007	Pass	0.049	Pass	0.009	Pass
Median House Price (log)	-0.019*	Pass	0.038***	Fail	-0.019*	Pass
Manufacturing share	-0.093	Pass	0.750	Pass	-0.108	Pass
Services share	0.559*	Pass	1.513	Pass	0.544*	Pass
Activity rate	-0.034	Pass	1.080*	Fail	-0.001	Pass
NVQ4	-0.143	Pass	1.091*	Pass	-0.132	Pass
Age 26-39 share	0.071	Pass	2.550***	Fail	0.116	Pass
Age 65+ share	-0.004	Pass	-2.190***	Fail	-0.068	Pass

Note: Table reports results from panel regression with fixed effects on local authority data for the period 1999-2016. A separate treatment effect δ_M has been estimated for the cities of Birmingham, Manchester, Liverpool, Edinburgh and Glasgow (equation 3). The final two columns pertain to estimates based on equation 2 excluding the above major cities from the sample. The regression includes local authority fixed effects and controls for time trends at the region level. The regression also includes time varying local authority variables. The major economic variables controls include for the share of the population with a degree level equivalent qualification (NVQ4), public sector GVA shares, manufacturing GVA shares, service sector GVA shares, the female economic activity rate, the overall economic activity rate and the pre-treatment local authority population. The demographic and sectoral variable controls include the pre-treatment local authority population the log of public sector GVA and the female activity rate. Standard errors are clustered at the local authority level and are also used for the parallel trend test. Region is at the NUTS1 level. Rejection of the parallel trend test implies a failure of the common trend assumption. See Appendix A4, Tables A.4.5, A.4.6, A.4.7, A.4.8, A.4.9, A.4.10, A.4.11 and A.4.12 for more details. * = p < 0.1, ** = p< 0.05, *** = p < 0.01

Table 3 repeats the analysis in Table 2 including the separate estimates (equation 3) for any additional average treatment effect δ_M for the major cities of Birmingham, Manchester, Liverpool, Edinburgh and Glasgow in column 3. When we attempted to estimate a separate additional treatment effect for the major cities, the point estimates for the general treatment effect on all local authorities δ_{LA} (column 2) for GVA, employment and population become smaller and less certain, such that we cannot be sure they are different from zero. The negative effect on house prices, however, is now larger and significantly different from zero. The estimate for the additional effects on the major cities (column 4) are positive and significant in the case of GVA and house prices. However, table 3 also reports that the parallel trend test fails[34] for the major cities (column 5) for GVA, population and house prices. This indicates the cities pre-treatment growth rates for these variables was significantly different (higher) in the pre-upgrade period and distorting the estimates of δ_M and δ_{LA}. Specifically, the estimates of δ_M is likely to be reflecting the continuation of faster pre-existing growth in major city populations, this is further discussed in section 4.2.2 around Figure 11.

[34] See appendix Table A.4.7.

A similar story plays out for the demographic and sectoral variables, whereby the general estimated treatment effects δ_{LA} are not estimated to be significantly different from zero except for the services share of GVA. The additional effect estimated for major cities is significant for the activity rate, NVQ4 share and both age shares. However, this is accompanied again by failed parallel trend tests for the activity rate and two age shares distorting the estimates of δ_M and δ_{LA}.

The fact that 6 out of 10 parallel trend tests fail for the major cities in column 5 of Table 3 suggest a failure of the common trend assumption for these cities (i.e., these cities were diverging from other[35] local authorities for the economic variables of interest prior to the upgrade completion beyond what are regression controls and fixed effects can account for). Therefore, under this approach, the DiD estimates for samples including these cities are likely invalid in terms of causality (including the results in Table 2) and we should focus on estimates that exclude[36] the major cities from the sample entirely. These estimates are reported in Table 3 (column 6), further noting that in column (7) all the parallel trend tests pass when we exclude these cities. While the point estimates are similar in sign and magnitude to the significant point estimates in Table 2, based on the estimated standard errors of the coefficients we cannot be sure they are not zero with the exception of house prices and the services share of GVA.

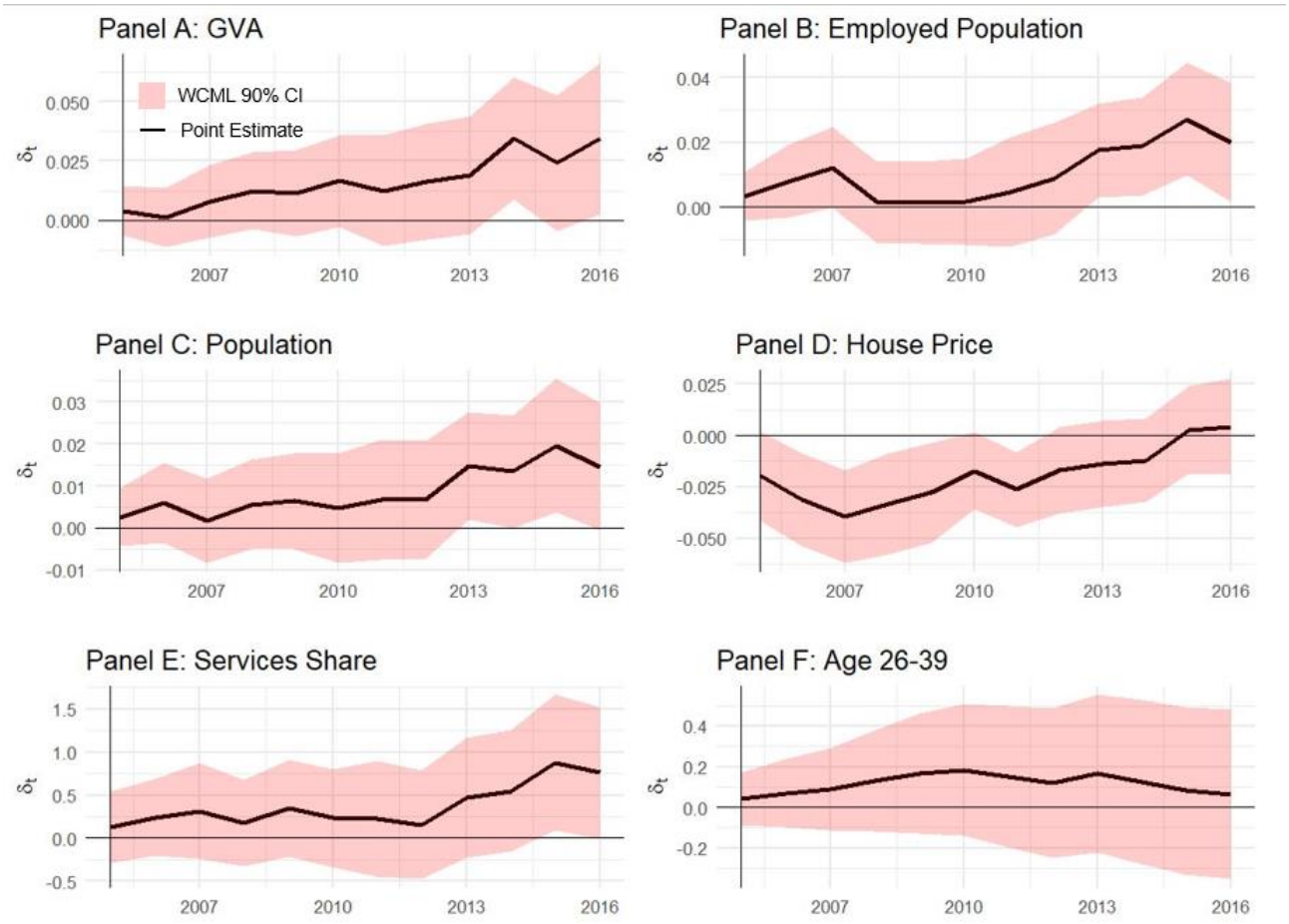
4.1.4 Average Treatment Effect Over Time - Excluding Major Cities

Figure 9 repeats the exercise of Figure 8 to estimate time varying treatment effects δ_t when excluding the major cities from the sample entirely. The analysis of Figure 9 is similar to Figure 8 in that the point estimates for GVA, employment, population and the services share are initially estimated as positive but not significantly different from zero, and then over time become larger and significantly greater than zero towards the end of the sample. Also as in Figure 8, we see the negative house price coefficient in Table 3 is driven by the initial period leading up to the financial crisis before fading and becomes negligible by the end of the sample. Taking the results from Table 3 and Figure 9 together indicates some tentative evidence of an economic impact of the WCML emerging over time.

[35] Other here means all local authorities other than the major cities identified on the WCML.

[36] Further work should look to explore and attempt to more rigorously control for the interaction between local authority size and time trends. This may enable analysis to be taken forward that includes these major cities.

Figure 9: Treatment effect δ_t for WCML by Year Excl. Major Cities



Note: Figure shows the estimates and 90 percent confidence intervals (red shaded area) for the estimates of δ_t exlcuding the major cities of Birmingham, Manchester, Liverpool, Edinburgh and Glasgow. Standard errors are clustered at the local authority level. Panels A-D are a log difference and Panels E and F are a difference in percentage point share.

4.2 Robustness

4.2.1 Average Treatment Effect - ECML

As a robustness exercise we conducted a placebo test whereby we assumed that it was the ECML stations that received an upgrade in 2005 and not the WCML stations. In practice, this means we set the treatment status to 0 for all WCML local authorities and to 1 for all ECML local authorities after 2004. We then repeated[37] the analysis in section 4.1.3. This test is motivated by the fact that the ECML stations present a natural control group to compare with the WCML stations (Table 1) and by the fact that there was no significant line wide improvement in GJT along the ECML that coincided in timing with the WCML upgrade (Figure 3).

[37] We omit repeating the analysis presented in section 4.1.1 and 4.1.2 as the results in section 4.1 have directed us to focus on the results excluding major cities. As such we begin by testing the impact of major cities on the ECML and presenting results without those major cities.

Table 4: Economic Impact (DiD) – ECML Placebo Incl. Major Cities

Dependent Variable	Effect (δ_{LA})	Parallel Trend (Pass/Fail)	Effect - Major Cities (δ_M)	Parallel Trend -Major Cities (Pass/Fail)	Effect - Excl. Major cities (δ)	Parallel Trend -Excl. Major Cities (Pass/Fail)
GVA (log)	0.005	Pass	0.036*	Fail	0.004	Pass
Employment (log)	0.015**	Pass	-0.007	Pass	0.015**	Pass
Population (log)	0.018**	Pass	-0.003	Fail	0.018**	Pass
Median House Price (log)	-0.020**	Pass	-0.06***	Fail	-0.020**	Pass
Manufacturing share	0.382	Pass	-0.397	Pass	0.403	Pass
Services share	-0.667	Pass	3.966***	Fail	-0.697	Pass
Activity rate	-0.262	Pass	0.608	Pass	-0.247	Pass
NVQ4	0.288	Pass	-0.169	Pass	0.277	Pass
Age 26-39 share	-0.139	Pass	2.154***	Fail	-0.149	Pass
Age 65+ share	-0.118	Pass	-2.017***	Fail	-0.123	Pass

Note: Table reports results from panel regression with fixed effects on local authority data for the period 1999-2016. A separate treatment effect δ_M has been estimated for the cities of Edinburgh, Leeds and Newcastle (equation 3). The final two columns pertain to estimates excluding the above major cities from the sample. The regression also includes time varying local authority variables. The major economic variables controls include for the share of the population with a degree level equivalent qualification (NVQ4), public sector GVA shares, manufacturing GVA shares, service sector GVA shares, the female economic activity rate, the overall economic activity rate and the pre-treatment local authority population. The demographic and sectoral variable controls include the pre-treatment local authority population the log of public sector GVA and the female activity rate. Standard errors are clustered at the local authority level and are also used for the parallel trend test. Region is at the NUTS1 level. Rejection of the parallel trend test implies a failure of the common trend assumption. See Appendix A4, Tables A.4.13, A.4.14, A.4.15, A.4.16, A.4.17, A.4.18, A.4.19 and A.4.20 for more details. * = p < 0.1, ** = p< 0.05, *** = p < 0.01

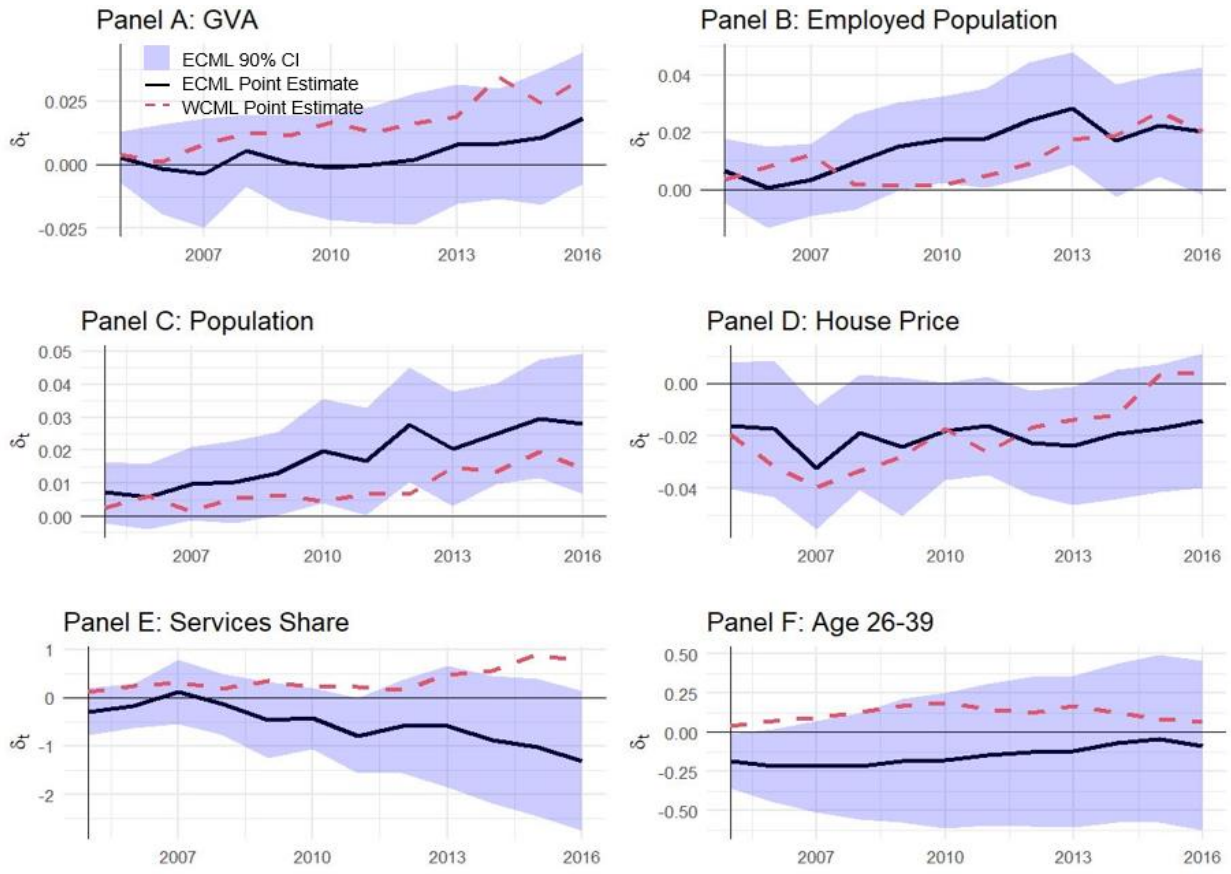
Given the analysis and tentative conclusions from section 4.1, we started by replicating the results of Table 3 under the placebo test. Table 4 reports these results with the additional effects estimated for the major cities on the ECML of Leeds, Newcastle and Edinburgh. As in Table 3, we see that the major cities are estimated to be diverging in trend in the pre-treatment period for a majority of the variables of interest and are thus likely to be distorting the estimates reported in column 2 and column 4 of Table 4 for δ_{LA} and δ_M. Of further note is the different behaviour of the major cities on the ECML compared to WCML. House prices are more negative in those major cities, whereas they are more positive in the WCML major cities. The point estimates for population and employment are negative, not positive as in the case of the WCML, though in each case they are not significant. This suggests different regional dynamics between major cities on the two lines. Putting this together, we again turn our focus to column 6, which reports the estimates for the ECML excluding those major cities. This shows significant positive point estimates for employment (+1.5%) and population (+1.8%), as well as a significant negative estimate for house prices (-2%). Overall, the results for the economic impacts in the ECML placebo test seem to be more[38] positive than for the WCML suggesting a lack of an identified effect of the WCML upgrade.

[38] This was also the case when we used the ECML as a direct control group for a DiD regression. That is, we estimated equation (1) using only local authorities on the ECML and WCML and year fixed effects instead of a year-region fixed effects. See appendix A3 for more details.

4.2.2 Average Treatment Effect Over Time – ECML vs. WCML

As in Figure 9, Figure 10 breaks out the treatment effect δ_t by year for the ECML placebo test (black line) and also includes the point estimates for the WCML (red dashed line) from Figure 9. Figure 10 is key to the conclusion of this report and shows that the red WCML line largely lies within the blue confidence interval of the ECML placebo test estimates indicating a lack of significant difference between the two estimates. The one exception is the effect on the share of the services sector of GVA, which is estimated to be larger for the WCML. The similarity of the profiles of the black and red dashed line, and the fact that we cannot find a significant difference in point estimates between the WCML exercise and ECML placebo exercise, creates considerable doubt over whether any of the findings in section 4.1 can be directly attributed to the WCML upgrade.

Figure 10: Treatment effect δ_t by Year (ECML Placebo excl. Major Cities)

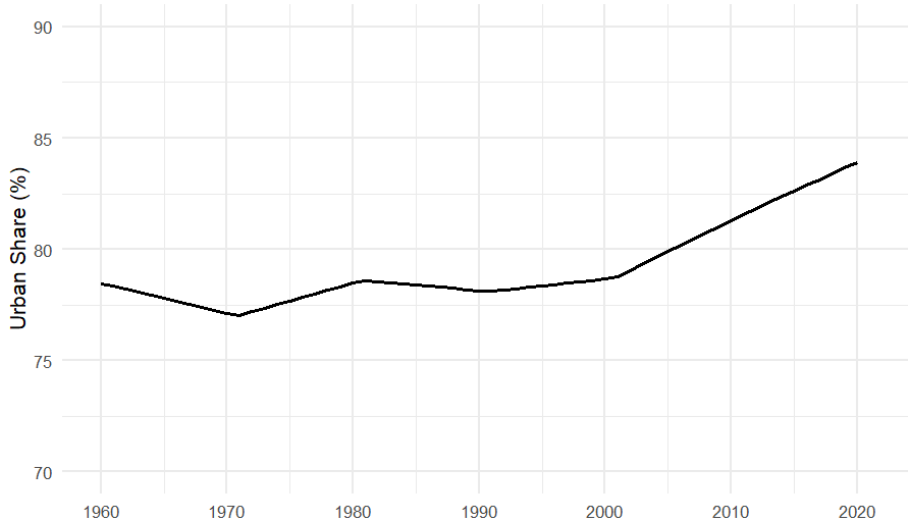


Note: Figure shows the estimates and 90 percent confidence intervals (blue shaded area) for the estimates of δ_t for the ECML mainline local authorities excluding Edinburgh, Leeds and Newcastle (equation 3). Standard errors are clustered at the local authority level. The red dashed line is the point estimate for the WCML as in Figure 9. Panels A-D are a log difference and Panels E and F are a difference in percentage point share.

However, while the analysis has not been able to clearly attribute economic impacts to the WCML upgrade, it does not necessarily mean that the WCML upgrade did not contribute to those economic impacts. A plausible explanation is an existing trend of employment and a younger population gravitating towards well connected urban areas, including local authorities along the WCML and ECML. By ensuring the WCML remained relatively well connected the WCML upgrade may have facilitated or prolonged this trend beyond what would have been likely in the absence of the upgrades. Another way of saying this is that the upgrade may have helped enable the two

regions/lines to stay on a common trend[39], where in absence of the upgrade it might have been expected that the WCML regions would have fallen behind the common trend of other regions e.g. below the blue swathes in Figure 10. Figure 11 shows how the urban population has increased in recent years in the UK beginning in the years prior to the upgrade. Given the robust growth in passenger volumes relative to the ECML (Figure 4), the WCML may have played a role in allowing WCML local authorities to take advantage of these broader trends and stay competitive with the rest of the country. We know that the upgrade was in part a catch up on foregone maintenance and improvements. So had the upgrade been unsuccessful, the significant point estimates recorded in Table 3 and Figure 9 may have been smaller.

Figure 11: Share of Population Living in Urban Areas in the UK



Source: ONS via World Bank. Access at: https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS?locations=GB

[39] This would represent a failure of the common trend assumption by expecting the WCML to not follow the ECML's trend in the post upgrade period, absent the upgrade.

5. Conclusion

This analysis sought to evaluate certain economic impacts of the WCML upgrade that were completed between 2005-2009. We compare the outcomes of local authorities with WCML stations to those without WCML stations and quantify the difference in a panel DiD setting using a set of area and time fixed effects. We found that local authorities exposed to the upgrade experienced a positive difference in GVA, employment and population following the upgrade. We also found a small, but statistically significant, difference on the percentage point (pp) share of young adults residing in local authorities with WCML stations and a difference on the share of output in the services sector. We found no significant difference in house prices and a limited impact on other sectoral and demographic characteristics.

In terms of attributing causality to the WCML upgrade, further analysis casts significant doubt as to the extent these estimates have identified an effect of the upgrade with the possible exception of the services sector share of GVA. Analysis breaking out the difference between major cities and other local authorities along the WCML indicates the significant point estimates reported were distorted by existing divergent pre-upgrade trends in the economic variables of major cities along the line. Furthermore, a robustness exercise using ECML local authorities found local authorities along the ECML experienced similar economic outcomes following the WCML upgrade.

The fact that the chosen methodology and data available for this analysis were unable to identify economic impacts resulting from the WCML upgrade does not necessarily mean the WCML upgrade did not produce them. A plausible conclusion, consistent with the success of the line in terms of passenger volumes, could be that the upgrade provided an improved foundation for the WCML regions to take advantage of pre-existing trends of population and economic activity shifting to urban areas, and in particular major cities. That is, without these important incremental upgrades the WCML local authorities may have compared less favourably to similar areas such as local authorities along the ECML, instead of our implicit assumption that they would, all else equal, have experienced similar outcomes without the intervention.

Finally, this work used relatively crude binary treatment definitions (local authority with WCML station). Future work could improve upon this by more rigorously defining exposure to the WCML upgrade by utilising improvements in journey times by local authority or more bespoke regions built up from more granular spatial sources. This approach might produce more precise estimates of the economic impacts. Other, albeit more black box, approaches specifically designed to overcome diverging pre-treatment tends (such as the synthetic control method[40]) may also be of benefit.

[40] Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program." Journal of the American statistical Association 105.490 (2010): 493-505.

Appendix

A1 Background

Upgrade

Table A.1.1: WCML upgrade – journey time changes by destination (to/from London Euston)[2]

Route (London Euston to/from)	Pre-project implementation	Post-project implementation	Journey Time Change
Birmingham	1 hr 43 mins	1 hr 30 mins	-13 mins
Coventry	1 hr 11 mins	1 hr 04 mins	-7 mins
Manchester	2 hr 36 mins	2 hr 06 mins	-30 mins
Liverpool	2 hr 53 mins	2 hr 30 mins	-23 mins
Crewe	2 hr 08 mins	1 hr 45 mins	-23 mins
Carlisle	4 hr 04 mins	3 hr 44 mins	-20 mins
Glasgow	5 hr 35 mins	5 hr 00 mins	-35 mins

Interviews

Table A.1.2

Name	Date of interview	Current role	Role during WCML upgrade
Neil Fleming	17/03/2022	Senior Economist - Expert Advisor, Rail Analysis, Department for Transport	Business case appraisal
Matt Dillon	15/07/2021	Associate Director, Arup	WCML Project Sponsor, Department for Transport (2003-2008)

Local Authority Locations

Table A.1.3

ECML	WCML	
Bassetlaw	Aylesbury Vale	North Lanarkshire
City of Edinburgh	Birmingham	North Warwickshire
County Durham	Bolton	Northampton
Darlington	Carlisle*	Nuneaton and Bedworth
Doncaster	Central Bedfordshire	Preston*
East Lothian	Cheshire East	Rugby
Hambleton	Cheshire West and Chester	Sandwell
Hertsmere	Chorley*	Solihull
Huntingdonshire	City of Edinburgh*	South Lakeland*
Leeds	Coventry	South Lanarkshire*
Newark and Sherwood	Dacorum	South Ribble*
Newcastle upon Tyne	Daventry	South Staffordshire
North Hertfordshire	Dudley	Stafford
Northumberland	Dumfries and Galloway*	Stockport
Peterborough	Eden*	Stoke-on-Trent
South Kesteven	Glasgow City*	Tamworth
Stevenage	Halton	Three Rivers
Wakefield	Lancaster*	Warrington*
Welwyn Hatfield	Lichfield	Watford
York	Liverpool	Wigan*
	Manchester	Wolverhampton
	Milton Keynes	

Note: Local authorities assigned if have a station on the main line within it's boundaries.
* Binary treatment begins in 2006 instead of 2005.
We have excluded the following local authorities from the analysis: City of London, Isles of Scilly Na h-Eileanan Siar, Shetland Islands, Orkney Islands and the Northern Ireland.

A2 Common Trend Test

We employ a common trend test for our variables of interest for the pre-treatment period. This test uses data from the pre-treatment period (1999-2004) to test the null hypothesis that the treatment and control group have the same time trend. This is achieved by running the following regression for each variable of interest y:

$$y_{i,t} = \alpha_i + \theta_{t,r} + \beta x_{i,t} + \lambda\, t\, 1_{WMCL} + \epsilon_{i,t}$$

Where α_i is a local authority fixed effect, $\theta_{t,r}$is a regional (NUTS1[41]) time fixed effect, $x_{i,t}$ are the same controls as used in the relevant DiD regression. The coefficient λ tests for any difference in linear time (t) trend in the pre-treatment period for the WCML local authorities relative to the other local authorities. The null hypothesis is that this is zero. If the estimated coefficient and standard errors leads us to reject that null hypothesis the results from the DiD are likely to be spurious as the common trend assumption is unlikely to be valid given the groups were not following a common trend in the pre-treatment period. For the specification in equation 3 that includes an effect for major cities we run the following specification:

$$y_{i,t} = \alpha_i + \theta_{t,r} + \beta x_{i,t} + \lambda\, t\, 1_{WMCL} + \lambda_M\, t\, 1_{WMCL\,Maj} + \epsilon_{i,t}$$

which augments the first specification by testing for any additional trend differences for major cities on the WCML (Birmingham, Manchester, Liverpool, Glasgow and Edinburgh). Any difference would be picked up by a significant estimate for λ_M.

[41] North East, North West, Yorkshire and the Humber, East Midlands, West Midlands, East of England, Greater London, South East, South West, Wales, Scotland. See: https://www.ons.gov.uk/methodology/geography/ukgeographies/eurostat (Accessed: 06/08/2022)

A3 Direct ECML Comparison - Regression Outputs

Table A.3.1 Economic impact (DiD) - for direct ECML comparison with time and local authority fixed effects. Standard errors reported in parentheses. * = p < 0.01, ** = p< 0.05, *** = p < 0.01. Feeds into Table 1.

	GVA (log)	No. employed by res. (log)	Population (log)	Med. House prices (log)
Intercept	-0.481***	0.117***	0.078***	2.070***
	(0.18)	(0.014)	(0.012)	(0.019)
Treatment	**-0.007**	**-0.010**	**-0.016***	**-0.005**
	(0.013)	**(0.009)**	**(0.009)**	**(0.015)**
Pre-treatment population (log)	0.716***	0.993***	0.990***	0.841***
	(0.012)	(0.008)	(0.008)	(0.011)
GVA from services (%)	0.005***	-0.003**	-0.002	-0.007***
	(0.002)	(0.001)	(0.001)	(0.002)
GVA from public sector (%)	-0.008**	-0.003*	-0.002	-0.004*
	(0.003)	(0.002)	(0.002)	(0.002)
GVA from manufacturing (%)	0.003	-0.001	-0.001	-0.004*
	(0.002)	(0.001)	(0.001)	(0.002)
Economic inactivity rate (%)	-0.001	-0.013***	0.003*	-0.003**
	(0.002)	(0.002)	(0.001)	(0.001)
NVQ4 (%)	0	0.002***	0.001***	0.001
	(0.001)	(0.001)	(0.000)	(0.001)
Female activity rate (%)	-0.000	0.000	0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
R-squared	0.996	0.997	0.998	0.985
R-squared Adj.	0.996	0.997	0.998	0.983
N	1094	1094	1094	1094

Table A.3.2 Demographic and Sectoral impact (DiD) - for direct ECML comparison with time and local authority fixed effects. Standard errors reported in parentheses. * = p < 0.01, ** = p< 0.05, *** = p < 0.01. Feeds into Table 1.

	Manufacturing Share (%)	**Services Share (%)**	**NVQ4 (%)**	**Economic activity rate (%)**	**Age 26 - 39 (%)**	**Age 65+ (%)**
Intercept	2.140	-3.020*	3.692	5.788***	1.667**	2.269*
	(1.654)	(1.592)	(2.599)	(1.080)	(0.816)	(1.348)
Treatment	**-0.206**	**1.141**	**-0.483**	**-0.273**	**-0.081**	**0.719**
	(0.453)	**(0.693)**	**(1.256)**	**(0.383)**	**(0.514)**	**(0.647)**
Pre-treatment population (log)	2.591***	1.746*	0.275	1.526**	0.930*	2.233***
	(0.925)	(0.894)	(1.370)	(0.579)	(0.468)	(0.776)
Female activity rate (%)	-0.051*	0.005	0.183***	0.649***	0.039**	-0.046**
	(0.025)	(0.021)	(0.043)	(0.020)	(0.018)	(0.020)
GVA from public sector (log)	-1.720	-0.425	0.761	0.689	1.113	-2.180*
	(1.535)	(1.477)	(2.301)	(0.961)	(0.776)	(1.299)
R-squared	0.945	0.964	0.864	0.914	0.930	0.898
R-squared Adj.	0.932	0.956	0.833	0.895	0.914	0.874
N	1098	1098	1094	1098	1044	1044

Table A.3.3 Economic impact Parallel Trend test - for ECML comparison with time and local authority fixed effects. Standard errors reported in parentheses. * = p < 0.01, ** = p< 0.05, *** = p < 0.01. Feeds into Table 1.

	GVA (log)	**No. employed by res. (log)**	**Population (log)**	**Med. House prices (log)**
Intercept	-0.481***	0.105	0.076	3.106***
	(0.126)	(0.112)	(0.100)	(1.128)
Parallel trend	**0.000**	**-0.000**	**0.000**	**-0.013**
	(0.004)	**(0.004)**	**(0.003)**	**(0.010)**
Pre-treatment population (log)	0.667*	1.018***	0.948***	1.801**
	0.394	0.352	0.323	0.760
GVA from services (%)	0.004	0.004	0.003	-0.001
	(0.003)	(0.004)	(0.003)	(0.007)
GVA from public sector (%)	-0.003	-0.002	-0.003	-0.002
	(0.004)	(0.003)	(0.002)	(0.006)
GVA from manufacturing (%)	0.006**	0.003	0.002	0.004
	(0.003)	(0.002)	(0.002)	(0.006)
Economic inactivity rate (%)	-0.003**	-0.014***	0.003**	-0.003**
	(0.002)	(0.002)	(0.001)	(0.003)
NVQ4 (%)	-0.000	-0.001	-0.001	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Female activity rate (%)	-0.003**	0.000	0.000	-0.004*
	(0.001)	(0.001)	(0.001)	(0.002)
R-squared	0.999	0.999	0.999	0.992
R-squared Adj.	0.999	0.999	0.999	0.989
N	367	367	367	337

Table A.3.4 Demographic and Sectoral Impact Parallel Trend test - for direct ECML comparison with time and local authority fixed effects. Standard errors reported in parentheses. * = p < 0.01, ** = p< 0.05, *** = p < 0.01. Feeds into Table 1.

	Manufacturing Share (%)	**Services Share**	**NVQ4 (%)**	**Economic activity rate (%)**
Intercept	332.938	-76.827	553.900	237.303
	(274.509)	(192.646)	(452.141)	(184.889)
Parallel Trend	**-0.183**	**0.047**	**-0.212**	**-0.046**
	(0.131)	**(0.093)**	**(0.216)**	**(0.087)**
Pre-treatment population (log)	5.857	1.181	-8.645	-10.552*
	(4.252)	(3.919)	(10.872)	(5.304)
Female activity rate (%)	-0.009	-0.001	0.136	0.659***
	(0.029)	(0.016)	(0.082)	(0.035)
GVA from public sector (log)	-3.192	-1.635	1.541	1.280
	(2.135)	(1.575)	(3.484)	(1.237)
R-squared	0.970	0.987	0.819	0.956
R-squared Adj.	0.963	0.983	0.773	0.945
N	371	371	367	371

A4 Regression Outputs

WCML – All sample

Table A.4.1 Economic impact (DiD) with region-time and local authority fixed effects. Standard errors reported in parentheses. * = p < 0.1, ** = p< 0.05, *** = p < 0.01. Feeds into Table 2.

	GVA (log)	No. employed by res. (log)	Population (log)	Med. House prices (log)
Intercept	-2.452***	-0.003	-0.071***	4.352***
	(0.027)	(0.017)	(0.016)	(0.024)
Treatment	**0.021****	**0.016****	**0.013***	**-0.016**
	(0.01)	**(0.008)**	**(0.007)**	**(0.011)**
Pre-treatment population (log)	0.849***	0.984***	0.985***	0.604***
	(0.004)	(0.003)	(0.003)	(0.004)
GVA from services (%)	0.006***	0.000	0.000	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)
GVA from public sector (%)	-0.006***	-0.001*	-0.001	-0.003**
	(0.001)	(0.001)	(0.001)	(0.001)
GVA from manufacturing (%)	0.005***	-0.001	-0.001**	-0.004***
	(0.001)	(0.000)	(0.000)	(0.001)
Economic inactivity rate (%)	0.000	-0.011***	0.005***	-0.001***
	(0.001)	(0.000)	(0.000)	(0.001)
NVQ4 (%)	0.000	0.001***	0.001***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Female activity rate (%)	0.000	0.001***	0.001***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
R-squared	0.996	0.996	0.997	0.99
R-squared Adj.	0.996	0.996	0.997	0.99
N	6082	6082	6082	5562

Table A.4.2 Demographic and Sectoral impact (DiD) with region-time and local authority fixed effects. Standard errors reported in parentheses. * = p < 0.1, ** = p< 0.05, *** = p < 0.01 Feeds into table 2

	Manufacturing Share (%)	Services Share (%)	NVQ4 (%)	Economic activity rate (%)	Age 26 - 39 (%)	Age 65+ (%)
Intercept	2.435	-23.592***	6.792*	9.336***	0.279	11.092***
	(2.942)	(2.312)	(3.771)	(1.698)	(1.194)	(1.631)
Treatment	**0.000**	**0.748****	**-0.006**	**0.101**	**0.423***	**-0.307**
	(0.452)	**(0.334)**	**(0.454)**	**(0.192)**	**(0.23)**	**(0.265)**
Pre-treatment population (log)	4.151***	4.501***	0.185	1.624***	1.415***	0.812**
	(0.635)	(0.495)	(0.815)	(0.369)	(0.27)	(0.368)
Female activity rate (%)	-0.005	-0.011	0.098***	0.661***	0.025***	-0.032***
	(0.01)	(0.008)	(0.02)	(0.009)	(0.005)	(0.006)
GVA from public sector (log)	-5.154***	-2.619***	1.137	-0.161	0.622*	-0.899*
	(0.804)	(0.627)	(1.029)	(0.463)	(0.347)	(0.474)
R-squared	0.951	0.967	0.873	0.909	0.952	0.945
R-squared Adj.	0.947	0.964	0.862	0.901	0.947	0.94
N	6138	6138	6082	6138	5688	5688

Table A.4.3 Economic impact parallel trend test with region-time and local authority fixed effects. Standard errors reported in parentheses. * = p < 0.1, ** = p< 0.05, *** = p < 0.01. Feeds into Table 2

	GVA (log)	No. employed by res. (log)	Population (log)	Med. House prices (log)
Intercept	1.072	-0.313	0.17	8.034*
	(2.259)	(1.362)	(1.195)	(4.462)
Parallel Trend Test	**0.004**	**0.000**	**0.000**	**0.005**
	(0.002)	**(0.001)**	**(0.001)**	**(0.006)**
Pre-treatment population (log)	0.494**	1.012***	0.956***	0.232
	(0.229)	(0.138)	(0.121)	(0.461)
GVA from services (%)	0.001	0.000	0.000	-0.010***
	(0.002)	(0.001)	(0.001)	(0.003)
GVA from public sector (%)	-0.002	0.000	0.000	-0.009***
	(0.002)	(0.001)	(0.001)	(0.003)
GVA from manufacturing (%)	0.005***	0.001	0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.003)
Economic inactivity rate (%)	0.000	-0.011***	0.005***	-0.001***
	(0.001)	(0.000)	(0.000)	(0.001)
NVQ4 (%)	0.000	0.000	0.000	0
	0.000	0.000	0.000	(0.001)
Female activity rate (%)	0.000	0.002***	0.002***	-0.001
	(0.000)	(0.000)	(0.000)	(0.001)
R-squared	0.998	0.998	0.999	0.991
R-squared Adj.	0.998	0.997	0.998	0.988
N	2050	2050	2050	1873

Table A.4.4 Demographic and Sectoral impact parallel trend test with region-time and local authority fixed effects. Standard errors reported in parentheses. * = p < 0.1, ** = p< 0.05, *** = p < 0.01 Feeds into Table 2

	Manufacturing Share (%)	**Services Share (%)**	**NVQ4 (%)**	**Economic activity rate (%)**	**Age 26 - 39 (%)**	**Age 65+ (%)**
Intercept	-7.705	39.805**	109.183**	189.111***	4.524	31.407***
	(20.579)	(18.745)	(51.35)	(22.711)	(5.385)	(6.678)
Parallel Trend Test	**0.137***	**0.077**	**-0.171**	**0.007**	**0.04**	**-0.051***
	(0.083)	**(0.091)**	**(0.156)**	**(0.075)**	**(0.029)**	**(0.029)**
Population (log)	1.994	-0.9	-5.375	-14.635***	1.073**	-1.091**
	(1.659)	(1.202)	(4.29)	(1.867)	(0.432)	(0.514)
GVA from public sector (log)	-0.004	-0.007	0.102***	0.640***	0.008**	-0.006
	(0.012)	(0.008)	(0.031)	(0.015)	(0.003)	(0.004)
Female activity rate (%)	-1.730*	-5.148***	-2.884	-0.115	0.36	-0.57
	(0.915)	(0.992)	(2.005)	(0.809)	(0.287)	(0.371)
R-squared	0.985	0.988	0.893	0.951	0.989	0.988
R-squared Adj.	0.981	0.985	0.865	0.938	0.986	0.985
N	2101	2101	2050	2101	1941	1941

WCML – Including Additional Effect on Major Cities

Table A.4.5 Economic impact including major cities term with region-time and local authority fixed effects. Standard errors reported in parentheses. * = p < 0.1, ** = p< 0.05, *** = p < 0.01. Feeds into Table 3.

	GVA (log)	No. employed by res. (log)	Population (log)	Med. House prices (log)
Intercept	-2.448***	0.002	-0.067***	4.355***
	(0.027)	(0.018)	(0.017)	(0.024)
Treatment	**0.015**	**0.009**	**0.007**	**-0.019***
	(0.01)	**(0.007)**	**(0.006)**	**(0.011)**
Treatment in cities	**0.044***	**0.056**	**0.049**	**0.038*****
	(0.024)	(0.037)	(0.036)	(0.014)
Pre-treatment population (log)	0.849***	0.984***	0.984***	0.603***
	(0.004)	(0.003)	(0.002)	(0.004)
GVA from services (%)	0.006***	0	0	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)
GVA from public sector (%)	-0.006***	-0.001*	-0.001	-0.003**
	(0.001)	(0.001)	(0.001)	(0.001)
GVA from manufacturing (%)	0.005***	-0.001	-0.001**	-0.004***
	(0.001)	(0.000)	(0.000)	(0.001)
Economic inactivity rate (%)	0	-0.011***	0.005***	-0.001***
	(0.001)	(0.000)	(0.000)	(0.001)
NVQ4 (%)	0	0.001***	0.001***	0
	(0.000)	(0.000)	(0.000)	(0.000)
Female activity rate (%)	0	0.001***	0.001***	0
	(0.000)	(0.000)	(0.000)	(0.000)
R-squared	0.996	0.996	0.997	0.99
R-squared Adj.	0.996	0.996	0.997	0.99
N	6082	6082	6082	5562

Table A.4.6 Demographic and Sectoral impact (DiD) including major cities term with region-time and local authority fixed effects. Standard errors reported in parentheses. * = p < 0.1, ** = p< 0.05, *** = p < 0.01. Feeds into Table 3

	Manufacturing Share (%)	**Services Share (%)**	**NVQ4 (%)**	**Economic activity rate (%)**	**Age 26 - 39 (%)**	**Age 65+ (%)**
Intercept	2.553	-23.354***	6.964*	9.505***	0.698	10.732***
	2.953	2.315	3.774	1.690	1.166	1.602
Treatment	**-0.93**	**0.559***	**-0.143**	**-0.034**	**0.071**	**-0.004**
	0.482	**0.312**	**0.498**	**0.190**	**0.182**	**0.245**
Treatment in cities	**0.750**	**1.513**	**1.091***	**1.080***	**2.550*****	**-2.190*****
	0.601	**1.063**	**0.618**	**0.561**	**0.617**	**0.643**
Pre-treatment population (log)	4.142***	4.484***	0.172	1.611***	1.382***	0.840**
	(0.635)	(0.495)	(0.815)	(0.369)	(0.27)	(0.368)
Female activity rate (%)	-0.006	-0.013	0.097***	0.660***	0.023***	-0.030***
	(0.01)	(0.008)	(0.02)	(0.009)	(0.004)	(0.005)
GVA from public sector (log)	-5.152***	-2.616***	1.140	-0.159	0.633*	-0.908*
	(0.804)	(0.625)	(1.028)	(0.460)	(0.338)	(0.466)
R-squared	0.951	0.967	0.873	0.909	0.954	0.947
R-squared Adj.	0.947	0.964	0.862	0.901	0.950	0.942
N	6138	6138	6082	6138	5688	5688

Table A.4.7 Parallel Trend Test for Economic impact including major cities with region-time and local authority fixed effects. Standard errors reported in parentheses. * = p < 0.1, ** = p< 0.05, *** = p < 0.01. Feeds into Table 3.

	GVA (log)	No. employed by res. (log)	Population (log)	Med. House prices (log)
Intercept	6.022*	4.564*	2.529	13.915***
	(3.0689)	(2.704)	(2.605)	(5.079)
Parallel trend	**0.002**	**-0.001**	**-0.000**	**0.004**
	(0.002)	**(0.001)**	**(0.001)**	**(0.006)**
Parallel Trend Major Cities	**0.010****	**0.010****	**0.005**	**0.014***
	(0.005)	**(0.004)**	**(0.004)**	**(0.008)**
Pre-treatment population (log)	0.059	0.582**	0.748***	-0.285
	(0.290)	(0.250)	(0.240)	(0.502)
GVA from services (%)	0.002	-0.001	-0.000	-0.012***
	(0.002)	(0.001)	(0.001)	(0.003)
GVA from public sector (%)	-0.003	0.000	-0.000	-0.011***
	(0.002)	(0.001)	(0.001)	(0.003)
GVA from manufacturing (%)	0.005***	0.000	0.001	-0.003
	(0.001)	(0.001)	(0.001)	(0.002)
Economic inactivity rate (%)	-0.001	-0.011***	0.004***	-0.003**
	(0.001)	(0.001)	(0.001)	(0.001)
NVQ4 (%)	0.000	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.001)
Female activity rate (%)	-0.001	0.002***	0.002**	-0.001
	(0.000)	(0.000)	(0.000)	(0.001)
R-squared	0.998	0.998	0.999	0.990
R-squared Adj.	0.998	0.997	0.998	0.987
N	1995	1995	1995	1823

Table A.4.8 Parallel Trend Test for Demographic and sectoral impact including major cities term with region-time and local authority fixed effects. Standard errors reported in parentheses. * = p < 0.1, ** = p< 0.05, *** = p < 0.01. Feeds into Table 3.

	Manufacturing Share (%)	**Services Share (%)**	**NVQ4 (%)**	**Economic activity rate (%)**	**Age 26 - 39 (%)**	**Age 65+ (%)**
Intercept	243.713**	305.344	-158.708	160.845*	159.934***	-166.301***
	(112.275)	(262.972)	(235.397)	(88.408)	(56.860)	(34.080)
Parallel Trend	**0.118**	**0.010**	**-0.169**	**-0.043**	**0.010**	**-0.021**
	(0.083)	**(0.063)**	**(0.164)**	**(0.075)**	**(0.026)**	**(0.026)**
Parallel Trend Major Cities	**0.200**	**0.540**	**0.021**	**0.319****	**0.240****	**-0.252*****
	(0.169)	**(0.423)**	**(0.397)**	**(0.135)**	**(0.093)**	**(0.055)**
Pre-treatment population (log)	-20.296*	-24.718	19.362	-11.752	-12.824**	16.482***
	(10.593)	(24.082)	(22.002)	(8.473)	(5.178)	(3.135)
Female activity rate (%)	-0.011	-0.008	0.094***	0.662***	0.005*	-0.001
	(0.011)	(0.007)	(0.030)	(0.015)	(0.003)	(0.003)
GVA from public sector (log)	-2.221**	-4.276***	-4.015**	-0.041	0.191	-0.110
	(0.886)	(0.542)	(1.916)	(0.862)	(0.233)	(0.236)
R-squared	0.985	0.989	0.884	0.947	0.989	0.992
R-squared Adj.	0.981	0.987	0.856	0.935	0.987	0.990
N	2046	2046	1995	2046	1896	1896

WCML – Excluding Major Cities

Table A.4.9 Economic impact excluding major cities (DiD) with region-time and local authority fixed effects. Standard errors reported in parentheses. * = p < 0.01, ** = p< 0.05, *** = p < 0.01. Feeds into Table 3.

	GVA (log)	No. employed by res. (log)	Population (log)	Med. House prices (log)
Intercept	-2.362***	-0.02	-0.075***	4.124***
	(0.025)	(0.016)	(0.014)	(0.021)
Treatment	0.016	0.01	0.009	-0.019*
	(0.011)	(0.007)	(0.006)	(0.011)
Pre-treatment population (log)	0.841***	0.984***	0.984***	0.625***
	(0.004)	(0.003)	(0.002)	(0.004)
GVA from services (%)	0.006***	0	0	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)
GVA from public sector (%)	-0.006***	-0.001*	0	-0.003**
	(0.001)	(0.001)	(0.001)	(0.001)
GVA from manufacturing (%)	0.005***	-0.001	-0.001**	-0.004***
	(0.001)	(0.000)	(0.000)	(0.001)
Economic inactivity rate (%)	0	-0.011***	0.005***	-0.001***
	(0.001)	(0.000)	(0.000)	(0.001)
NVQ4 (%)	0	0.001***	0.000***	0
	(0.000)	(0.000)	(0.000)	(0.000)
Female activity rate (%)	0	0.002***	0.001***	0
	(0.000)	(0.000)	(0.000)	(0.000)
R-squared	0.996	0.996	0.997	0.99
R-squared Adj.	0.995	0.996	0.997	0.99
N	5992	5992	5992	5508

Table A.4.10 Demographic and Sectoral impact excluding major cities (DiD) with region-time and local authority fixed effects. Standard errors reported in parentheses. * = p < 0.1, ** = p< 0.05, *** = p < 0.01. Feeds into Table 3.

	Manufacturing Share (%)	Services Share (%)	NVQ4 (%)	Economic activity rate (%)	Age 26 - 39 (%)	Age 65+ (%)
Intercept	1.586	-21.102***	5.909*	8.198***	-0.219	10.686***
	(2.769)	(2.171)	(3.56)	(1.585)	(1.072)	(1.477)
Treatment	-0.108	0.544*	-0.132	-0.001	0.116	-0.068
	(0.488)	(0.313)	(0.501)	(0.19)	(0.18)	(0.238)
Pre-treatment population (log)	4.232***	4.282***	0.304	1.746***	1.528***	0.771**
	(0.623)	(0.484)	(0.802)	(0.359)	(0.253)	(0.347)
Female activity rate (%)	-0.006	-0.013	0.096***	0.658***	0.021***	-0.028***
	(0.01)	(0.008)	(0.02)	(0.009)	(0.004)	(0.005)
GVA from public sector (log)	-5.153***	-2.614***	1.073	-0.179	0.521	-0.776*
	(0.811)	(0.631)	(1.04)	(0.463)	(0.336)	(0.462)
R-squared	0.951	0.966	0.871	0.908	0.956	0.948
R-squared Adj.	0.946	0.963	0.859	0.9	0.952	0.944
N	6048	6048	5992	6048	5598	5598

Table A.4.11 Economic impact excluding major cities parallel trend test with region-time and local authority fixed effects. Standard errors reported in parentheses. * = p < 0.01, ** = p< 0.05, *** = p < 0.01. Feeds into Table 3.

	GVA (log)	No. employed by res. (log)	Population (log)	Med. House prices (log)
Intercept	-1.073	-0.313	0.175	6.127**
	(0.001)	(0.001)	(0.683)	(3.092)
Parallel Trend Test	**0.002**	**-0.001**	**0.000**	**0.004**
	(-0.002)	**(0.001)**	**(0.001)**	**(0.006)**
Pre-treatment population (log)	**0.705*** **	**1.062*** **	**0.992*** **	**0.416**
	(0.142)	**(0.081)**	**(0.075)**	**(0.337)**
GVA from services (%)	0	-0.001	-0.001	-0.011***
	(0.002)	(0.001)	(0.001)	(0.003)
GVA from public sector (%)	-0.002	0	0.000	-0.009***
	(0.002)	(0.001)	(0.001)	(0.003)
GVA from manufacturing (%)	0.005***	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Economic inactivity rate (%)	-0.001	-0.011***	0.004***	-0.002*
	(0.001)	(0.001)	(0.001)	(0.001)
NVQ4 (%)	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.001)
Female activity rate (%)	0.000	0.002***	0.002***	-0.001
	(0.000)	(0.000)	(0.000)	(0.000)
R-squared	0.998	0.998	0.998	0.991
R-squared Adj.	0.998	0.997	0.998	0.988
N	2015	2015	2015	1850

Table A.4.12 Demographic and Sectoral impact excluding major cities parallel trend test with region-time and local authority fixed effects. Standard errors reported in parentheses. * = p < 0.01, ** = p< 0.05, *** = p < 0.01. Feeds into Table 3.

	Manufacturing Share (%)	Services Share (%)	NVQ4 (%)	Economic activity rate (%)	Age 26 - 39 (%)	Age 65+ (%)
Intercept	-5.757	54.485***	112.962**	197.375***	8.077*	28.366***
	(21.259)	(15.081)	(52.36)	(22.264)	(4.703)	(6.429)
Parallel Trend Test	**0.113**	**0.003**	**-0.17**	**-0.035**	**0.01**	**-0.02**
	(0.085)	**(0.062)**	**(0.165)**	**(0.075)**	**(0.026)**	**(0.027)**
Population (log)	2.116	-1.453	-5.742	-14.946***	0.865**	-0.940*
	(1.683)	(1.124)	(4.356)	(1.854)	(0.388)	(0.495)
GVA from public sector (log)	-1.753*	-5.095***	-2.908	-0.09	0.389	-0.593
	(0.919)	(0.998)	(2.015)	(0.815)	(0.287)	(0.372)
Female activity rate (%)	-0.004	-0.008	0.100***	0.639***	0.007**	-0.006
	(0.012)	(0.008)	(0.031)	(0.015)	(0.003)	(0.004)
N	0.984	0.988	0.891	0.95	0.989	0.988
R-squared	0.981	0.985	0.863	0.937	0.986	0.985
R-squared Adj.	2066	2066	2015	2066	1906	1906

ECML Placebo Test Including Major Cities

Table A.4.13 Placebo Test for Economic impact including major cities term with region-time and local authority fixed effects. Standard errors reported in parentheses. * = p < 0.01, ** = p< 0.05, *** = p < 0.01. Feeds into Table 4.

	GVA (log)	No. employed by res. (log)	Population (log)	Med. House prices (log)
Intercept	-2.454***	-0.005	-0.071***	4.349***
	(0.027)	(0.017)	(0.016)	(0.023)
Treatment	**0.005**	**0.015****	**0.018****	**-0.020****
	(0.011)	**(0.007)**	**(0.007)**	**(0.010)**
Treatment in Major Cities	**0.036***	**-0.007**	**-0.003**	**-0.060*****
	(0.021)	**(0.021)**	**(0.017)**	**(0.012)**
Pre-treatment population (log)	0.849***	0.984***	0.985***	0.604***
	(0.004)	(0.003)	(0.002)	(0.004)
GVA from services (%)	0.006***	0.000	0.000	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)
GVA from public sector (%)	-0.006***	-0.001**	-0.001	-0.003**
	(0.001)	(0.001)	(0.001)	(0.001)
GVA from manufacturing (%)	0.005***	-0.001	-0.001**	-0.004***
	(0.001)	(0.000)	(0.000)	(0.001)
Economic inactivity rate (%)	-0.000	-0.011***	0.005***	-0.001***
	(0.001)	(0.000)	(0.000)	(0.001)
NVQ4 (%)	0.000	0.001***	0.001***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Female activity rate (%)	-0.000	0.001***	0.001***	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
R-squared	0.996	0.996	0.997	0.997
R-squared Adj.	0.996	0.996	0.997	0.990
N	6082	6082	6082	5562

Table A.4.14 Placebo Test for Demographic and Sectoral impact including major cities term with region-time and local authority fixed effects. Standard errors reported in parentheses. * = p < 0.01, ** = p< 0.05, *** = p < 0.01. Feeds into Table 4.

	Manufacturing Share (%)	**Services Share (%)**	**NVQ4 (%)**	**Economic activity rate (%)**	**Age 26 - 39 (%)**	**Age 65+ (%)**
Intercept	2.408	-23.329***	6.780*	9.379***	0.419	10.977***
	(2.942)	(2.324)	(3.768)	(1.693)	(1.192)	(1.619)
Treatment	**0.382**	**-0.667**	**0.288**	**-0.262**	**-0.139**	**-0.118**
	(0.405)	**(0.506)**	**(0.632)**	**(0.254)**	**(0.239)**	**(0.316)**
Treatment in Major Cities	**-0.397**	**3.966*** **	**-0.169**	**0.608**	**2.154*** **	**-2.017*** **
	(0.924)	**(1.263)**	**(0.699)**	**(0.458)**	**(0.431)**	**(0.339)**
Pre-treatment population (log)	4.154***	4.455***	0.188	1.616***	1.390***	0.826**
	(0.634)	(0.497)	(0.813)	(0.369)	(0.270)	(0.364)
Female activity rate (%)	-0.005	-0.012	0.0098***	0.660***	0.025***	-0.032***
	(0.010)	(.0008)	(0.019)	(0.009)	(0.005)	(0.006)
GVA from public sector (log)	-5.161***	-2.563***	1.130	-0.149	0.648*	-0.904*
	(0.802)	(0.630)	(1.025)	(0.463)	(0.347)	(0.470)
R-squared	0.951	0.967	0.873	0.909	0.952	0.946
R-squared Adj.	0.947	0.964	0.862	0.901	0.948	0.941
N	6138	6138	6082	6138	5688	5688

Table A.4.15 Parallel Trend Test for Placebo, Economic impact including major cities term with region-time and local authority fixed effects. Standard errors reported in parentheses. * = p < 0.01, ** = p< 0.05, *** = p < 0.01. Feeds into Table 4.

	GVA (log)	**No. employed by res. (log)**	**Population (log)**	**Med. House prices (log)**
Intercept	1.347	0.344	0.646*	11.226***
	(2.048)	(0.552)	(0.363)	(3.787)
Parallel Trend	**0.001**	**-0.002**	**-0.001**	**-0.002**
	(0.003)	**(0.002)**	**(0.002)**	**(0.006)**
Parallel Trend Major Cities	**0.014***	**0.004**	**0.004***	**0.038****
	(0.008)	**(0.003)**	**(0.002)**	**(0.017)**
Pre-treatment population (log)	0.498***	0.978***	0.932***	0.018
	(0.194)	(0.060)	(0.044)	(0.387)
GVA from services (%)	0.002	-0.001	-0.000	-0.013***
	(0.002)	(0.001)	(0.001)	(0.003)
GVA from public sector (%)	-0.003*	0.000	-0.000	-0.012***
	(0.002)	(0.001)	(0.001)	(0.003)
GVA from manufacturing (%)	0.005***	0.001	0.001	-0.003
	(0.001)	(0.001)	(0.001)	(0.002)
Economic inactivity rate (%)	-0.001	-0.012***	0.004***	-0.003**
	(0.001)	(0.001)	(0.001)	(0.001)
NVQ4 (%)	0.000	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.001)
Female activity rate (%)	-0.001	0.002***	0.002***	-0.001
	(0.000)	(0.000)	(0.000)	(0.001)
R-squared	0.998	0.998	0.999	0.990
R-squared Adj.	0.998	0.997	0.998	0.988
N	1995	1995	1995	1823

Table A.4.16 Parallel Trend Test for Placebo, Demographic and Sectoral impact including major cities term with region-time and local authority fixed effects. Standard errors reported in parentheses. * = p < 0.01, ** = p< 0.05, *** = p < 0.01. Feeds into Table 4.

	Manufacturing Share (%)	**Services Share (%)**	**NVQ4 (%)**	**Economic activity rate (%)**	**Age 26 - 39 (%)**	**Age 65+ (%)**
Intercept	7.412	67.651***	96.170	187.752***	16.965***	-29.924***
	(28.328)	(134.177)	(85.505)	(15.059)	(17.995)	(5.824)
Parallel Trend	**-0.010**	**-0.061**	**-0.011**	**0.053**	**-0.048**	**-0.018**
	(0.127)	**(0.066)**	**(0.173)**	**(0.087)**	**(0.032)**	**(0.029)**
Parallel Trend Cities	**0.044**	**1.096****	**-0.135**	**-0.104**	**0.220*****	**-0.130*****
	(0.167)	**(0.509)**	**(0.35)**	**(0.087)**	**(0.072)**	**(0.033)**
Pre-treatment population (log)	1.192	-18.321	7.914	1.898	-2.085	4.350***
	(4.311)	(12.478)	(8.55*9)	(2.191)	(1.762)	(0.735)
Female activity rate (%)	-0.010	-0.007	0.092***	0.662***	0.006**	-0.002
	(0.011)	(0.007)	(0.030)	(0.015)	(0.003)	(0.003)
GVA from public sector (log)	-2.169**	-4.432***	-4.068**	-0.023	0.158	-0.112
	(0.881)	(0.536)	(1.924)	(0.867)	(0.235)	(0.238)
R-squared	0.985	0.989	0.884	0.947	0.989	0.992
R-squared Adj.	0.981	.0987	0.856	0.935	0.986	0.990
N	2046	2046	1995	2046	1896	1896

ECML Placebo Test Excluding Major Cities

Table A.4.17 Placebo test excluding Major Cities Economic impact with region-time and local authority fixed effects. Standard errors reported in parentheses. * = p < 0.01, ** = p< 0.05, *** = p < 0.01. Feeds into Table 4.

	GVA (log)	No. employed by res. (log)	Population (log)	Med. House prices (log)
Intercept	-2.316***	-0.019	-0.073***	4.060***
	(-0.025)	(-0.015)	(-0.014)	(-0.02)
Treatment	**0.004**	**0.015****	**0.018****	**-0.020****
	(-0.011)	**(-0.007)**	**(-0.007)**	**(-0.01)**
Pre-treatment population (log)	**0.837*****	**0.984*****	**0.984*****	**0.631*****
	-0.004	**-0.003**	**-0.002**	**-0.004**
GVA from services (%)	0.006***	0	0	-0.006***
	(-0.001)	(-0.001)	(-0.001)	(-0.001)
GVA from public sector (%)	-0.006***	-0.001*	0	-0.003**
	(-0.001)	(-0.001)	(-0.001)	(-0.001)
GVA from manufacturing (%)	0.005***	-0.001	-0.001**	-0.004***
	(-0.001)	(0)	(0)	(-0.001)
Economic inactivity rate (%)	0	-0.011***	0.005***	-0.001***
	(-0.001)	(0)	(0)	(-0.001)
NVQ4 (%)	0	0.001***	0.000***	0
	(0)	(0)	(0)	(0)
Female activity rate (%)	0	0.002***	0.001***	0
	(0)	(0)	(0)	(0)
R-squared	0.996	0.996	0.997	0.991
R-squared Adj.	0.995	0.995	0.997	0.99
N	5956	5956	5956	5472

Table A.4.18 Placebo test excluding Major Cities, Demographic and Sectoral impact with Region-time and local authority fixed effects. Standard errors reported in parentheses. * = p < 0.01, ** = p< 0.05, *** = p < 0.01. Feeds into Table 4.

	Manufacturing Share (%)	**Services Share (%)**	**NVQ4 (%)**	**Economic activity rate (%)**	**Age 26 - 39 (%)**	**Age 65+ (%)**
Intercept	0.974	-19.599***	6.053*	8.055	-0.193	10.490***
	(2.719)	(2.146)	(3.489)	(1.559)	(1.052)	(1.441)
Treatment	**0.403**	**-0.697**	**0.277**	**-0.247**	**-0.149**	**-0.123**
	(0.405)	**(0.505)**	**(0.634)**	**(0.253)**	**(0.238)**	**(0.319)**
Pre-treatment population (log)	4.319***	4.086***	0.096***	0.658***	0.021***	-0.028***
	(0.6019)	(0.486)	(0.794)	(0.358)	(0.252)	(0.343)
Female activity rate (%)	-0.006	-0.014*	0.096***	0.658***	0.021***	-0.028***
	(0.010)	(0.008)	(0.020)	(0.009)	(0.004)	(0.005)
GVA from public sector (log)	-5.223***	-2.479***	1.061	-0.168	0.553	-0.776*
	(0.811)	(0.638)	(1.037)	(0.466)	(0.338)	(0.460)
R-squared	0.951	0.966	0.871	0.908	0.956	0.949
R-squared Adj.	0.946	0.963	0.859	0.900	0.952	0.944
N	6012	6012	5956	6012	5562	5562

Table A.4.19 Parallel Trend Test for Placebo test excluding Major Cities, Economic impact with Region-time and local authority fixed effects. Standard errors reported in parentheses. * = p < 0.01, ** = p< 0.05, *** = p < 0.01. Feeds into Table 4.

	GVA (log)	No. employed by res. (log)	Population (log)	Med. House prices (log)
Intercept	-2.230***	-0.161	-0.159	3.693***
	(0.176)	(0.133)	(0.120)	(1.062)
Parallel Trend	**0.000**	**-0.002**	**-0.001**	**-0.002**
	(0.003)	**(0.002)**	**(0.002)**	**(0.006)**
Pre-treatment population (log)	0.829***	1.022***	1.005***	0.699***
	(0.051)	(0.039)	(0.034)	(0.171)
GVA from services (%)	-0.000	-0.001	-0.001	-0.011***
	(0.002)	(0.001)	(0.001)	(0.003)
GVA from public sector (%)	-0.002	0.000	-0.000	-0.009***
	(0.002)	(0.001)	(0.001)	(0.003)
GVA from manufacturing (%)	0.005***	0.001	0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.003)
Economic inactivity rate (%)	-0.000	-0.011***	0.004***	-0.002*
	(0.001)	(0.001)	(0.001)	(0.001)
NVQ4 (%)	0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.001)
Female activity rate (%)	-0.000	0.002***	0.002***	-0.001
	(0.000)	(0.000)	(0.000)	(0.001)
R-squared	0.998	0.998	0.998	0.991
R-squared Adj.	0.997	0.997	0.998	0.988
N	2003	2003	2003	1838

Table A.4.20 Parallel Trend Test for Placebo test excluding Major Cities, Demographic and Sectoral impact with Region-time and local authority fixed effects. Standard errors reported in parentheses. * = p < 0.01, ** = p< 0.05, *** = p < 0.01. Feeds into Table 4.

	Manufacturing Share (%)	**Services Share (%)**	**NVQ4 (%)**	**Economic activity rate (%)**	**Age 26 - 39 (%)**	**Age 65+ (%)**
Intercept	7.412	67.651***	96.170	187.752***	16.965***	30.975***
	(25.151)	(17.595)	(58.612)	(24.608)	(7.347)	(8.368)
Parallel Trend	**-0.001**	**-0.082**	**-0.018**	**0.039**	**-0.048**	**-0.021**
	(0.128)	**(0.067)**	**(0.177)**	**(0.084)**	**(0.031)**	**(0.030)**
Pre-treatment population (log)	2.129	-1.525	-5.758	.14.948***	0.835**	-0.930*
	(1.680)	(1.124)	(4.379)	(1.855)	(0.389)	(0.479)
Female activity rate (%)	-0.003	-0.008	0.098***	0.639***	0.008**	-0.006
	(0.012)	(0.008)	(0.031)	(0.015)	(0.003)	(0.004)
GVA from public sector (log)	-1.735*	-5.175***	-2.961	-0.093	0.364	-0.594
	(0.916)	(0.997)	(2.026)	(0.821)	(0.289)	(0.373)
R-squared	0.984	0.988	0.891	0.950	0.989	0.988
R-squared Adj.	0.980	0.986	0.863	0.937	0.987	0.985
N	2054	2054	2003	2054	1894	1894