Critical realist perspectives on the urban growth system

David Waite
University of Glasgow, UK

Abstract
Why some cities have been able to sustain economic growth more than others, and how particular technological or sectoral break points shift a city’s economic performance favourably (or unfavourably), present ongoing conundrums. Whilst single factor accounts – such as the skilled city and the creative city – have the appeal of analytical parsimony, economic geographers suggest that a multiplicity of structures, processes and events typically sit behind how growth originates and is subsequently shaped. Given such complexity, how can we develop appropriate approaches to theorising causality within urban economic systems? This paper presents the case that critical realism may perform a useful ground-clearing role. With a layered ontology at its core, coupled with recent literature providing greater guidance for empirical application, it is argued that critical realism may present a complementary explanatory perspective.

Keywords
Critical realism, urban growth, mechanisms

Introduction
Processes of urban economic growth and change present many questions, from the nature and dimensions of growth processes themselves to considerations of how growth may be managed and governed. For a number of authors, the conundrum of explaining urban economic growth, remains stubbornly present (Huggins, 2016), whilst related issues of resilience and adaptation prompt us to consider why growth ‘may flourish in some [contexts], but wither on the vine in others’ (Storper, 2010: 2042). Such issues reflect a central point of enquiry for many economic geographers concerned with uneven development, however, there are persistent challenges relating to the identification of appropriate scales of analysis and causal dynamics (Overman, 2003; Sunley, 2008). The contribution of this paper is to give a line of sight for how the critical realist approach may aid explanatory work on urban growth processes; indeed, this paper aims to set out procedures.
for empirical application, though, as will be shown, a full guidebook remains elusive (Lewis, 2011). The key to understanding the role of critical realism, it is argued, is a recognition of the more tentative and provisional truth claims that result. However, promise is suggested in terms of the ability of this meta-theoretical approach to begin to identify the multiple causal processes likely supporting growth.

There have been numerous points of inflection between economic geographers and critical realism. In the domain of ‘geographical political economy’, critical realism is presented as an almost foundational setting for theorising causality: ‘informed by critical realism, the geographical political-economic approach advanced here copes with context, diversity and variety by recognising that causal powers inherent in structures can be realised only contingently in particular time-space contexts’ (Pike et al., 2016: 127). For others, the positioning of researcher and research object, required by critical realism, holds appeal (MacKinnon, 2010: 30). Alternatively, Jones and Murphy (2010) turned to critical realism as a means to justify a ‘practice’ orientation in economic geography. This contribution hinges on the somewhat surprising intervention that critical realism can be meshed with actor-network theorising, which is typically conceived as a flat ontology. Adding to the literature, Pratt (1995) sought to explore the analytical potential in terms of assessing the dynamics of an industrial estate; others have explored critical realism’s distinctions relative to Marxism (Roberts, 2001); while path dependency has also been framed in a critical realist manner (Martin and Sunley, 2006; cf Yeung, 2019b). One of the key recurring challenges is how to put critical realism into practice (Pratt, 1995) and economic geographers are not alone in confronting this issue (Davidsen, 2009; Larsen and Lindkvist, 2014). Whilst recent contributions attempt to set out wayfinders for empirical work (Price and Martin, 2018), earlier applied attempts confronted numerous ambiguities and methodological grey areas (Yeung, 1997). Therefore, whilst critical realism is not foreign to economic geographers – and offers a distinguishing approach to geographical economics (Mäki and Marchionni, 2009; Sayer, 2004) – neither is it an easy bedfellow. This is the case, despite Peck’s (2012) plea that economic geographers could fruitfully engage with developments in heterodox economics, where critical realism remains prominent (Mäki and Marchionni, 2009; Lee, 2002). Economic geographers have flirted with critical realism, in essence, yet thorough, applied engagements remain scarce.

Before sketching out where, and in what form a critical realist perspective may make a contribution to research on urban economic growth, some basic propositions central to critical realism are important to lay out. Critical realists, at the most simple level, take the view that there are surface processes configured by underlying causal mechanisms which operate in an observer-independent manner. As Harding and Blokland (2014: 18) remark, critical realism hinges on a layered ontology, where ‘on the surface there is the empirical domain of observable experiences. At a deeper level is the actual domain consisting of events, whether or not they are observed. At the deepest level is the real domain that is made up of structures and mechanisms that cause events’ (also see Castellacci (2006:861)). Critically, such causes are distinct from what we observe and are not immediately reducible from events; they have emergent properties, in other words (Elder-Vass, 2005; Sayer, 2013). It is the aim of critical realist research to posit the underlying causal mechanisms and suggest what role they play in a social system (Sayer, 2000). Critical realism has and is developing a diverse heritage, and it is fair to reflect that, while centred on core stances or perspectives – the core ontological roots of critical realism as a meta-theory (Hastings, 2021) – authors from different disciplines engage with critical realism in varying ways through specific, empirically focused theoretical developments (or what Davidsen (2009: 52) refers to as ‘domain specific’ approaches).

Open systems are a critical feature of a realist ontology. Chick and Dow (2005: 366) remark on the nature of open systems that they are ‘not atomistic… structure and agency are interdependent… boundaries around and within the social or economic system are mutable… and identifiable social structures are embedded in larger structures’ (also see Derbyshire, 2020). Rather than committing to
a view of permanent regularities in social processes – ‘if event x then event y’ – the open systems view, for critical realists, posits that such conjunctions are, when apparent, space-time limited, Lawson, 1997). This view of open systems is an ontological one and is not to be confused with economic geographers’ concern for assessing intersecting territorial and relational processes (this concerns the varying spatial interactions and reaches of geographic phenomena, reflecting a research scope issue). Also central to the critical realist approach are the notions of the transitive and intransitive domains, which refer to the domains of knowledge and to the domains of being respectively (Sayer, 2000). A notable feature of the critical realist approach is to try and avoid collapsing one into the other; for economic geographers, this may be akin to saying that the urban economy changes because our theories of it require it to change. Sayer (2000: 10–11) remarks here: ‘When theories change (transitive dimension) it does not mean that what they are about (intransitive dimension) necessarily changes too: there is no reason to believe that the shift from a flat earth theory to a round earth theory was accompanied by a change in the shape of the earth itself’. A related point concerns the extent to which a priori theoretical stances can and should guide empirical work, and how apriorism frames abstraction (Isaksen 2016; Van Meeteren et al., 2016). Sayer (2000: 3) talks of and rejects the apriorism of neoclassical economics, because of the attachment to strict mechanisms for understanding economic phenomena. He notes elsewhere (Sayer, 2004: 1783) of a ‘cavalier’ attitude to abstraction in this line of research, where claims to ‘proceed by the power of our assumptions’ appears to excuse ‘checking whether the forms of abstraction were safe’ in the first place. This is a plea for both exercising caution in suggesting mechanisms and critically appraising the manner in which we let theories guide us toward these. Indeed, this aligns with Lawson’s foundational work in Economics and Reality (1997) which emphasises the need to take the objects and phenomena of economies seriously.

In looking to confront the problem of explaining urban growth through a critical realist approach, some delineation of the empirical object in question is warranted initially. Here, a recent flurry of outputs on urban theory offers some core markers (Scott, 2021; Scott and Storper, 2015; Storper and Scott, 2016; Walker, 2016). In this literature, a key feature of the debate concerns the wedge between a claim for certain universal characteristics of the urban and a recognition of and emphasis on particularity. Here, Scott and Storper (2014: 8) propose the idea of an ‘urban land nexus’ and agglomeration as clarifying what the urban must exhibit; these are proposed as essential characteristics of the urban (regardless of the differential manner in which these may emerge in different contexts). For others, however, this approach glides over the differential nature of the urban, particularly as observed by scholars concerned with the Global South (Robinson, 2011: 19), who seek to ‘unsettle parochially derived theoretical certainties’, while for Mould (2016: 168) there is a wariness about conceiving the urban as an ‘economic technological system’. A further front of debate concerns a turn to assemblage theory by some, with Storper and Scott (2016: 1126) suggesting that this approach reflects a ‘descriptive, anecdotal and notably indiscriminate approach to urban investigation’, lacking ‘theoretical guideposts’ and thus falling short on providing an analytical means to determine ‘significant relationships or to distinguish between the trivial and the important’ (also see Walker, 2016: 177). Notably, for Storper and Scott (2016: 1128) their critique of assemblage-inspired approaches turns to signal the merits of critical realism: ‘Critical realism has long provided a way out of this kind of dead-end by insisting on the importance of necessary relationships, causal powers and theoretical abstraction as fundamental to the identification of the central properties and conditions of existence of social phenomena’. Though Walker (2016: 176) differs on other points from Scott and Storper (2014)2, a common ground is hinted at in terms of the potential of critical realism: ‘Deep investigation of complex cases, sorting out the wheat from the chaff, is a perfectly valid scientific approach, no less powerful than broad comparisons across wide fields of data. Both are subject to the same demand for theory-guidance in making sense of evidence. The best guide to method in the
social sciences is still Sayer’. At play in the aforementioned debates and cleavages concerning urban theory are core issues concerning what is inherently constitutive of the urban versus what merely impacts the urban, and how, in grappling with the inherent complexity of urban areas as empirical objects, analysts can make claims as to what is critical as opposed to what is trivial.

The paper proceeds, in the next section, to set out the nature of urban economies and the challenges of explaining how they grow and change. This leans on the economic geography literature most obviously, though key contributions in urban economics are also presented. From this point, the paper turns to consider a prominent example in the literature that has sought to present a causal account of an urban economic growth problem. In pointing out potential limitations in this example, the next section then goes on to sketch out what a critical realist mode of reasoning might imply for analysing urban economic growth (this is set out with a view to application, rather than providing a fully worked through account, given the constraints of space). The paper ends with reflections on the promise and problems of a critical realist approach (giving reference to Yeung’s (2019a; 2019b, 2021) and Gong and Hassink’s (2020) recent contributions to the discussion in economic geography).

The nature of urban economies and the growth system

The urban growth problem reflects the challenges in specifying why a city exhibits a particular growth path, and why this may differ from other cities. Moreover, the problem is concerned with the break points and events that may shape trajectories of growth. Economic growth is different from but potentially supportive of economic development (Feldman et al., 2016). The latter takes a wider concern for human well-being and distributional processes within the economy. More narrowly, growth refers to a change in production outputs (gross value added), and in their recent work on diverging urban trends across UK cities, Martin et al. (2018) refer to ‘GVA per employed worker’ and ‘output growth’ (also see Martin et al., 2016; Tyler et al., 2017). It is this narrow perspective on growth that underpins the focus of this paper. Of course, research design also shapes the perspective we take on urban growth. Indeed, how we look at the output trajectories of two cities over time, will differ, methodologically, from how we uncover a particular event that has been seen to shape a specific industry trajectory at a single city.

Explaining urban economic growth is not a new challenge. Looking back to previous exposures of urban change, such as that given by Chinitz (1961) – who compared Pittsburgh and New York, in what the author described as ‘two immersions in regional economics’ (Chinitz, 1961: 279) – we can see an exemplar of rich case expositions where with the context of place and its entrepreneurship, capital, labour and land interdependencies come to the fore. Chinitiz’s viewpoint on urban economies remains potent (Glaeser, 2009). Richardson’s (1972) survey of the literature on urban growth, meanwhile, drew key links between growth and location theory, while concerns for city size took centre stage in other work (Richardson, 1973). Other influences on understanding urban economic change rest, notably, in Jane Jacobs’ work, who presented anecdotes formed into broader ‘conjectures’ about how cities work (Duranton, 2017: 1872), and despite being considered ‘foremost a journalist and an opinionated citizen’ by Duranton (2017: 1872), her ideas still leaves notable traces in the literature.

Indeed, one area where Jacobs’ work still resonates is in the literature on agglomeration economies, which is a common approach for framing urban economic growth. Agglomeration typically refers to spatially concentrated interactions, which working at scale or density and leading to cumulative development dynamics, may generate an uplift in productivity of between 3% and 8% (Puga, 2010: 207). The idea that firms and workers are recursively drawn to, and beneficially interact within, large or dense urban areas is a core notion of agglomeration. Different dimensions, urbanisation or localisation, reflect the sectoral depth or breadth of economic activities undertaken, whilst three external economies (sources) – input-output; labour market depth; knowledge – along with
three underpinning mechanisms – sharing; matching and learning – have a central place in the literature in terms of conceptualising how agglomeration comes about (Duranton and Puga, 2003; Puga, 2010). Offering contextualisation, geographers have stressed the distinctions between routine and uncertainty-dominated activities (Scott et al., 2002) and the varied spatial settings agglomeration processes exhibit (Phelps and Ozawa, 2003). The agglomeration literature suffers, in terms of explanation, from ‘Marshallian equivalence’ (Overman and Puga, 2008: 1) which means that it is difficult to distinguish, empirically, which of the sources is most prominent within a spatial concentration (cf Diodato et al., 2018). Progress has been made, however, in distinguishing sorting processes from agglomeration effects, so as to more accurately ascertain the economic benefits directly derived from agglomeration. Indeed, it is useful to distinguish when density and scale induce productivity from when existing productivity induces density and scale (Gaubert, 2018). Further challenges rest in setting out the sizes necessary for agglomeration economies to take hold; whilst some, such as the World Bank (2009), suggest a bigger is better logic, Frick and Rodríguez-Pose (2017) show that, subject to varying contexts, small to mid-size cities appear to exhibit the most favourable growth trajectories (perhaps due to the minimisation of diseconomies). In general, and whilst urban and regional economic literatures are replete with empirical cases highlighting the advantageous processes that can be found in certain urban contexts, the identification of causal processes remains a challenge. This is a limitation: we can espouse a general story of certain cities exhibiting productivity advantages, yet we struggle to discern the underpinning dynamics. The agglomeration literature has little to say, moreover, on how processes of cumulative causation are triggered (Storper, 2009, 2010, 2011a).

Understanding urban economic growth trajectories require researchers to grapple with the interaction of ‘structures’, ‘processes’ and ‘events’, it has been argued (Storper, 2011a: 344). The challenge rests on both explaining the initial cause of growth and then what sustains and cultivates a growth trajectory over time (where subsequent shocks and events may also have a bearing). Storper (2010, 2013) further contests that central to improving our understanding of urban economic growth is making progress on a chicken and egg problem; this refers to whether urban growth is set off by people following firms, or firms following people. The literature divides on this point (Hoogstra et al., 2017). Glaeser (2011) sides with the latter proposition – as reflected in a long-running concern for amenities (Glaeser et al., 2001) – however, for Storper (2013: 27), where the rise of the sun belt in the United States warrants explanation, a more plausible argument is that the link works the other way (the movement of manufacturing to take advantage of new policy instruments led firms to areas with warmer climates, people then followed). Here aim is also taken at NNEUE (new neo-classical urban economics) approaches more broadly, which hinge on the spatial interactions between firms, workers and developers. Equilibrium outcomes stem from these interdependencies (Glaeser, 2007, 2008) and trade-offs are pervasive (e.g., high wages are met with high prices) (Glaeser, 2008: 4). The model rests on positions of ‘indifference’ held by firms (hiring more workers), workers (choosing locations) and developers (building more houses) (Glaeser and Gottlieb, 2009: 984) leading to classes of cities offering different packages of wages, amenities and housing provisions. The assumptions here present issues as to whether they are realistic (Storper, 2011a; also see Mäki, 2018) and, more particularly, whether the simultaneity required by the equilibria elide key causal moves (Storper, 2010: 2031). In seeking to advance causal understandings, alternatively, Storper (2010) points to the interactions between specialisation, human capital and institutions. This is perhaps an uncontroversial triad, and if you place developers and the land market as place-based institutions, these factors would also overlap with the components in Glaeser’s notion of spatial equilibrium (2008). However, how the triad is seen to interact, fundamentally differs.

First, specialisation relates to the growth of a particular industry or activity at a place through potentially cumulative effects. However, there are challenges in explaining, one, where the ‘seeds’ for specialised growth are sewn, and, two, where they are successfully cultivated.
Moreover, how places respond to external shocks – do they see precipitous employment decline or can they upgrade to higher-value activities – is a key consideration. Further issues relate to whether a city will see favourable economic dynamics if it is sectorally specialised or diverse, and how the evolution of firms (and product life cycles) respond to a city’s specialisation (Duranton and Puga, 2001). For others, distinguishing absolute specialisation (pure size) from relative specialisation (intra-city share) (Kemeny and Storper, 2015), and functional from sectoral specialisation draws analytical focus (Duranton and Puga, 2005; Martin et al., 2016). Second, human capital perspectives are informed by literature which evidences that skilled people move to cities so they can be around other skilled people. Learning mechanisms and experience accumulation processes are well rehearsed here (de la Roca and Puga, 2017; Glaeser and Maré, 2001) and are coupled with new perspectives on how worker ability and confidence shape spatial sorting (de la Roca et al., 2014). Moreover, the agglomeration literature points to the matching mechanisms alleged to be at play in urban labour markets (Puga, 2010). Here particular competencies held by workers can be matched with the precise specialisations of firms that have cause to apply the competencies. Recent work by Balland et al. (2020), meanwhile, prompts us to consider the differential nature of knowledge and complexity in cities. Finally (third), institutions refer to rule-setting and policymaking types as well as the cultures and norms that colour a socio-economic context (Bradford and Bramwell, 2014; Rodríguez-Pose, 2020). Rodríguez-Pose (2013: 1041) has argued that a thorny matter concerns whether good institutions generate growth or follow from growth. Simply put, do we see strong performing economies with good institutions because the institutions have followed the growth, or vice versa? Additionally, the role of institutions has been found to be important – relative to factors of first-order geography – in shaping regional development (Ketterer and Rodríguez-Pose, 2018), whilst the scales at which formal institutions are effective and possess capacity have been elaborated on (Rodrigues-Pose and Ezcurra, 2011). In parts of spatial economics, institutions appear to be framed in terms of the regulation of land use and housing, or in terms of the role they play in mitigating diseconomies that may emerge from greater density and scale. Reading across these three pillars of the growth problem, the explanatory dilemma emerges given that each framing has its causal antecedents in each other (i.e., there is persistent endogeneity). Storper (2010) argues, more specifically, that explaining human capital as a growth driver may require reversion to specialisation and institutions. Likewise with specialisation, one may need to revert to forms of human capital and institutions in order to explain the emergence of a particular economic concentration. The same endogeneity issue emerges for institutions.

In framing the initial cause for an area, moreover, Storper (2009, 2013) is at pains to stress that initial first-round wins for an area – an initial and original growth trigger – may evolve over time, and as an industry or production process matures and differentiates, opportunities for other locations to plug into the industry may become apparent. This gives a broader spatial and temporal logic to urban economic growth, beyond the three factors stated prior. In this rendition, a growth impulse can emerge from the changing knowledge patterns and processes of production within sectors, presenting opportunities for other places to support and host sector activities at later points in time. From this inherently evolutionary conception, productive links may be made to allied areas of economic geography – such as related variety (Boschma and Iammarino, 2009) and path creation (Neffke et al., 2011) – which present further perspectives on urban and regional growth trajectories.

A prominent causal account

Given the array of different starting points in the literature (as outlined in prior paragraphs), coupled with their blind spots, generating more plausible explanatory accounts of urban economic growth remains a key challenge. Storper (2010, 2011a, 2013) attempts this by hunting for a middle ground
between the complex, idiosyncratic worlds embraced by geographers and the more restrictive but explanatory neat assumptions in spatial economics. In commenting on the challenges of explaining urban growth, Storper (2011a, 2011b: 343) has remarked that a ‘framework for the field’: ‘… should deal with both large-scale, continuous or “regular” growth processes, and be able to incorporate “shocks” and “events”. It should, moreover, incorporate endogenous regional dynamics that are potentially non-linear and that themselves generate further sudden change and local selection processes’. To do this, rich, granular data, covering multiple time periods, that can accommodate the ruptures and changes which trigger or sustain a particular growth trajectory, are required (Storper, 2011a: 344). Additionally, the researcher may usefully assemble data on events and shocks, such as the ‘lock in of localisations’, so one could estimate ‘… their effects jointly with structural determinants’ (Storper, 2011a: 344). These suggested ways forward are set out and incrementally elaborated on in multiple places (Storper, 2010, 2011a; Storper et al., 2015) yet operationalising such approaches warrant further consideration. At the heart of Storper’s suggestion is a move toward greater causal precision so the triggers of growth and subsequent growth cultivation processes can be comprehensively theorised; ‘though the challenge is daunting, only an approach that melds structure, events and processes, and hence can tackle directions of causality, is likely to advance us significantly …’ (Storper, 2011a: 344). Furthermore, in acknowledging the required links between the ‘on the ground sensibilities of geographers’ with the ‘large scale quantitative analyses favoured by economists’ accommodation for pluralism is hinted at (Storper, 2011a: 344; 2011b).

Such concerns form the backdrop of Storper et al.’s (2015) book – the Rise and Fall of Urban Economies – which interrogates the diverging growth trajectories of San Francisco and Los Angeles over four decades. As the reader moves toward the latter chapters of the book, the argument crystallises around the idea that the differing activities of business groups and elites are central to the divergent performance of the two cities; San Francisco had favourable business ecologies to support the new economy in simple terms. The structure of this rich book, which is set out in a ‘whodunit’ style (Storper et al., 2015: 26), allow the authors to ‘eliminate’ or show ‘to be inadequate’ alternative explanations for this divergence (in Aoyama et al. 2017: 155). The question may be posed, nevertheless, that even if it can be argued that the Bay Area Council supported a ‘zeitgeist’ without comparison in southern California (Storper et al., 2015), does this offer the definitive or dominant account of the diverging growth paths? Indeed, the ‘whodunit’ approach deployed in the Rise and Fall is ultimately revealing as a logic for explanation; by eliminating explanations one-by-one, identifying the prized causal factor appears to be the aim. Touching on the challenges of achieving this, Scott (Aoyama et al., 2017: 152) points to the continuing legacies of aerospace decline (in Los Angeles), and questions the ‘weight of emphasis placed on ‘progressive, forward-looking business culture[s]’ presented in the book. Additionally, Essletzbichler (2016: 1069) posits that the role of offshore manufacturing coupled with favourable tax approaches may be ‘equally or more important in accounting for rapidly rising median wages in the Bay Area’. Therefore, an analytical challenge in Storper et al. (2015) – and apparent in similar endeavours such as Rodrigues-Pose and Hardy (2021), who seek to explain the differing trajectories of Barcelona and Madrid – is the apparent quest to unearth the one, or at least the dominant, explanation that resolves the comparative question. This prompts the broader question – does this manner of reasoning (the whodunit) risk engendering a form of reductionism in the clamour for the definitive (or primary) causal account? Or do these accounts reflect an appropriate and reasonable balance of ‘parsimony and expansiveness’ given the empirical object in focus (Storper, 2011b: 14)?

Drawing on recent contributions to the management literature around configurational theorising, there is merit, I argue, in considering ‘how causal forces at different levels of analysis contribute in complex ways to outcomes of interest’ (Furnari et al., 2021: 791-792). Though some accounts of urban change may more obviously hinge on a single definitive cause, which clearly reflects the
trigger behind a sustained process of cumulative causation – for example, Moretti’s (2013) discussion on the divergent growth paths of Seattle and Albuquerque – other cases of growth and decline more likely reflect multiple causes taking effect (so understanding the interaction of multiple causes – including their layering and temporal sequencing – is the key challenge). The following section considers whether critical realist-inspired approaches may offer some guidance for grappling with the issue of causal multiplicity.

**Critical realist approaches**

In this section, considerations for applying critical realist approaches to questions related to urban growth are sketched out, and I explore the potential for critical realism to add explanatory depth. To be clear, what follows is not addressing the urban economic growth problem in terms of explaining why the growth trajectories of two cities have taken different tacks over time (as summarised in the prior section). Rather, the focus is on how a critical realist-inspired approach may support explanations of an important component part of a single city’s economic growth trajectory. This analytical focus aligns with Walker’s (2016: 176) remark, noted prior, that critical realism has merit for the ‘deep investigation of complex cases’ and parallels with Martin’s (2016: 1626) account of appreciative theorising, involving the grounded analyses of ‘the evolution of firms, industries and sectors, based on detailed, concrete causal accounts of specific case studies’ (also see Castellacci, 2006). With an allied focus furthermore, work by Grillitsch et al. (2022) seeks to give an account of regional economic development from an agent focus, whilst in considering the ‘dark matter’ of context and regional differences, ‘insufficient attention to the sources of action at the level of the actor’ was pointed to by Storper (2009: 17). Debates in heterodox economics about thick description and the consequences for theorising causality would also appear to link to these concerns (Downward et al., 2002; also see Yeung, 2019a: 228).

When applying the meta-theory of critical realism in social science applications (Fleetwood, 2017; Hoddy, 2019), there is merit, I argue, in engaging with Bhaskar’s legacy as much as Sayer has a foundational role, whilst the domain-specific guides provided by Lawson (1997) and Mäki (Mäki and Marchionni, 2009) *inter alia* provide further guidance. In this respect, there is a need to go beyond the core questions presented by Sayer (2000: 16) who sets out a series of perspectival questions – ‘What does the existence of this object/practice presuppose? What are its preconditions? Can/could object A, e.g., capitalism, exist without B, e.g., patriarchy? What is it about this object which enables it to do certain things?’ – to a more fully worked through methodology. That is not to dispute the pertinence of Sayer’s questions – they are the important ‘what is going on here’ starting points – but rather to ask how one would build an analysis around this. In cognate disciplines, scholars have made determined efforts recently to put into effect a critical realist approach, taking on Bhaskar’s steps (Fletcher, 2017; Hu, 2018; Steffansen, 2016). This hinges on deploying an RRREIC approach typically (Isaksen, 2016: 249–250), which has been summarised as follows:

Resolution (R1) of the phenomenon or event into its components, involving a multiplicity of causes that come together in a complex way to generate the phenomenon; Redescription (R2) of these component causes in an explanatorily significant way; Retroduction (R3) of these causes to antecedently existing events or states of affairs [I add retroduction in here as with Mingers and Standing (2017)]; Elimination of alternative explanations; Identification of the generative causes; and Correction of earlier findings in the light of this analysis. (Steffansen, 2016: 125)

It is this approach that we will now begin to sketch out. We condense RRREIC into three steps, however, for ease of use – description/explication; retroduction/retrodiction; correction/refinement – and this broadly accords with Hu (2018) and Harding and Blokland (2014: 18).
To give sight lines for application, reference is given to empirical work involving the collection of interview data from 23 organisations, made up of multi-location firms and related industry bodies, in Edinburgh’s (UK) financial services sector (between 2012 and 2013) (Waite, 2015). The research examined how the ‘firm-territory nexus’ (Dicken and Malmberg, 2001) in Edinburgh is shaped by the activities of multi-location firms, and how the emergence, anchoring and consolidation of the activities undertaken by such firms in the city hinges on globally orchestrated firm processes and routines, meshing with conditions and capacities particular to the territory of Edinburgh (as a second-tier financial centre) (Waite, 2015). The link to urban economic growth rests on what we know about the local economy. To give the current position broadly - Edinburgh (the city) exhibits strong economic performance; the financial and insurance sector accounts for a relatively high share of jobs in the city (9.5% of all jobs compared with 3.3% across Scotland and 3.5% across Great Britain), the sector contributes an estimated £5.4 billion to the city economy (22% of the city’s total GVA) and is linked to areas of growth such as Fintech (a summary of data is at City of Edinburgh Council 2021; also see Hutton and Shalchi, 2021). Therefore, a case can be made that shifts in this sector are an important consideration for understanding the city’s economic performance. Moreover, the literature on multinational enterprises (MNEs) and regional development, gives us strong reason to believe that these firms – and how they organise what they do where – are significant for shaping urban and regional economic growth trajectories (Iammarino and McCann, 2013).

The first steps are to analytically frame the phenomena of concern. As a starting point, demi-regularities can be seen to reflect ‘rough trends or broken patterns in empirical data’ (Fletcher, 2017: 185) or ‘partial, approximate, rough-and-ready regularities or patterns in the flux of events’ (Fleetwood, 2017:47). Such patterns then warrant a causal account – why do we see what we are seeing? Contrastive demi-reg (Lawson, 1997:206) are potent within economic geography, and can be seen to reflect some of the most fundamental questions within the area – San Francisco’s economy has grown at a faster rate than Los Angeles’ economy (Storper et al., 2015); cities in the North of the UK have tended to grow at a slower rate than cities in the south-east (Martin et al., 2016; Tyler et al., 2017). Additionally, why does the same tech industry evolve differently in different cities (Castellacci, 2006: 875)? The contrastive logic may also apply for a single urban area, such as considering the performance of different technology-focused industries in the same city. In the Edinburgh example, the demi-reg is Edinburgh’s strong productivity performance compared with other second-tier UK cities (see Centre for Cities, 2022) coupled with its disproportionate share of financial services activities. These observations set off the search for explanatory mechanisms.

Taking a particular research design choice to look at MNEs, given what we know about the local economy, there is then the challenge of marshalling the interview and accompanying documentary data. To do this an a priori framework of absorptive capacity was deployed to provide core dimensions to order the data collected – access (firm activities within reach), anchor (the fixity of activities within reach), diffuse (local spillovers from the fixed activities). The argument here – and moving absorptive capacity from its typical terrain of firms and knowledge – is that the apriorism acts as a way (through metaphor effectively) to categorise the data collected, and provides a plausible framework by which to think about firms in territories. This is not without risks, of course (when data are unduly forced into an a priori category). This refers to R1 and R2 in the RRREIC schema introduced prior.

There is then, in the next step, the question of what sits behind the firm-territory nexus that we can observe. What mechanisms are driving multi-locations firms to access and anchor in Edinburgh, and how does their anchoring shift (or not) into new roles (with a view to upgrading)? Unpicking what might drive or sit behind the existence of a demi-reg, critical realists use retroduction as a way to reveal causal processes; and this can be broadly or narrowly conceived. Retroduction, as Pratt
(2009: 380) notes in broad terms, involves ‘the movement between the real object, the necessary and contingent conditions, a rational abstraction, and the specification of causal mechanisms’, whilst Downward and Mearman (2007: 88) refer to ‘a thought operation that moves between knowledge of one thing to another, generating an explanation that embraces ontological depth’. While consisting of iterative analytical transitions between data to abstraction (and back again) (Sayer, 2013: 24), retroduction is widely seen to lack clear steps and specification (Fletcher, 2017; Lee, 2002: 793), and how particular mechanisms can be identified as necessary and prominent has been subject to much debate (Hodgson, 2004: 67; Isaksen, 2016). In the Edinburgh-financial services work – as determined through a coding process hinging on iterative analysis through the data collection process – the following mechanisms may be posited: network configuration, industry regulation, firm cost-control logic, financial centre reputation (similar to Walker’s (2016) ‘symbolic city’) and human capital (Waite, 2015). In essence, these proposed mechanisms appear to go a long way to explaining the firm-territory nexus and why we see multi-location financial services firms in Edinburgh undertaking the activities they do, both reflecting the lure of Edinburgh as well as the limits on what activities the city can appear to host (in terms of the latter, for example, network configuration reflects the tendency that only certain high-value activities will locate in London). (Waite, 2015). Moreover, relative changes in each of the mechanisms may amplify or suppress the operability of another mechanism (Blom and Moren, 2011).

However, the suite of explanatory possibilities does not give a sense of which mechanisms have prominence, or whether there are first-order or second-order mechanisms at play. This step requires retroduction. To make the distinction: retroduction is a creative process of specifying what mechanisms may explain the phenomena we observe; retroduction on the other hand switches the focus – if the mechanism was in play, what phenomena may we expect to see (Fleetwood, 2002: 38). This may be more simply framed as considering outcomes to causes (reduction) then causes to outcomes (retroduction) (McAvoy and Butler, 2018: 164). Where in retroduction you discover the mechanisms to explain a phenomenon, in retroduction the researcher looks at how already known mechanisms may be in effect to lead to a phenomenon (or empirical pattern). In applied work, both modes are likely to be observed somewhat in tandem (Mingers and Standing, 2017); we will know some mechanisms that likely hold up the firm-territory nexus, but others we will need to find and abstract from the data collected. In the Edinburgh-financial services example, retroduction – asking what outcomes the suggested mechanisms lead to – resulted in an ordering of the mechanisms across further categories – meta, mediating channels and direct. The meta refers to the network and industry-wide processes that determine what activities tend to go where in the financial services sector; mediating channels refer to the inherited and calculative processes both from the firm and the territory-side (so for a firm, designating a firm centre of excellence as part of corporate strategy; for the territory, promoting reputational qualities of the financial centre); and direct, whether firm activities are strengthened through a merger or acquisition or diminished through the loss of a key manager (so activities and transactions). This provides us with a conception of the firm-territory nexus taking form through mechanisms operating at different levels. At this point then, the a priori framework of access anchor and diffuse – as critical temporal components of the firm-territory nexus – are developed further with layers of abstraction that show how the phases or periods (access, anchor, diffuse) are shaped by different mechanisms working at different levels (the meta, mediating channels and direct) (Waite, 2015).

The final step concerns how we can eliminate and correct factors, and distinguish the necessary from the contingent; and this step in the RRREIC approach seems, for this author at least, to warrant further consideration. In terms of our concern for the urban growth system, the RRREIC approach offers the potential – through the retroductive and retrodictive steps – to sort which mechanisms appear to have more explanatory bites than others. However, the particular explanatory contribution of critical realism – in terms of the urban growth problem – rests on firmer ground in the exploratory
steps above, such as ‘inference to best explanation’ (and the retroductive and retrodictive steps) (Price and Martin, 2018: 95), than adjudicating more definitively on what mechanisms are the prime cause of an empirical pattern. This presents critical realism as a modest exercise of discovery (Hodgson, 2004; Melia 2020).

Some further remarks on the nature of causal elimination are warranted, however, as it is a feature of the later steps of the RRREIC approach presented above and a term used by Storper et al. (2015) to help justify the ‘whodunit’ approach. Ultimately the phenomena of the urban economic system lead to hesitancy in progressing this step because distinguishing necessary or central mechanisms is different from discarding a mechanism altogether because it is insufficient (Storper et al. in Aoyama et al., 2017: 155). Indeed, how might a set of factors interweave at a local place to shape a growth dynamic (Sayer, 2000: 15) and what is the empirical warrant for discarding a factor when it may plausibly interact with an allegedly dominant factor in an important way? A factor may be inadequate but vital, in other words. The analysis clearly needs to endeavour to sort out the key mechanisms, but on what grounds can we dispense with factors that may appear to be peripheral *prima facie*?

**Resurgent concerns**

The story of critical realism and economic geography can be characterised as one of fitful and passing engagement, as researchers periodically raise their heads above the parapet to reflect on and sometimes promote the meta-theory. Indeed, recent interest in critical realism can be evidenced in Yeung’s (2019a, 2019b, 2021) and Gong and Hassink’s (2020) contributions and these outputs raise interesting questions about attempts to add causal depth to geographic analyses. First, Yeung’s (2021) initiative to ‘bridge’ related variety with strategic coupling raises issues about a causal mechanism (strategic coupling) becoming available to the related variety literature in terms of theorising extra-local linkages. Second, Gong and Hassink’s (2020: 484) claim that related variety presents ‘fundamental philosophical problems’ and lacks an explicit mechanism for growth is perhaps interesting to consider further given the role that knowledge spillovers have been theorised to play (Content and Frenken, 2016). Together, these papers demonstrate both the continued draw to critical realism, and the challenges of applying this perspective to growth-related concerns.

In this paper, an approach has been set out that serves to grapple with explanatory mechanisms for a particular urban economic growth context. In doing this – and with critical realism posed as an explanatory underlabourer – the advantages are sensitivity to context and a way to move past pure description by suggesting the joint operation of mechanisms. Economic growth in a particular urban context is likely to be driven by multiple causal processes, and the explanatory steps of reduction and retrodiction may provide some explanatory headway. However, the critical realist vantage is a limited one, and more gains are made in the exploratory steps, it has been argued, than in the move to arrive at definitive mechanistic accounts.

Ultimately, the appetite for heterodox economics and allied approaches such as critical realism will hinge on existing analytical preferences and theoretical commitments. For some, establishing definitive mechanisms underpinning urban economic phenomena, within existing frameworks and tractable models (Mäki, 2018; Mäki and Marchionni, 2009), may reflect the key steps to be taken. For others who adopt a form of pluralism (Barnes and Sheppard, 2010; Martin, 2021) – where diversity of viewpoints is broadly valued rather than reconciliation and conceptual tidiness per se – critical realist-inspired approaches may find favour. From such a pluralist stance, it is not a case of one approach being more valid, but – when hinging on a ‘common ground for discourse’ (Dow, 2004: 282) or a ‘common empirical referent’ (Martin, 2021: 22) as reflected by the urban growth problem – drawing merit from different approaches and the various aspects of urban phenomena they shine a light on (Martin, 2018; Pike, 2019; also see Grüne-Yanoff and Marchionni, 2018).
Explaining urban economic growth remains a great challenge. If starting points for explanation remain many and varied – as befits the complexity of the issue – this paper contests that critical realist-inspired approaches can fruitfully contribute alongside other approaches (Marchionni, 2004). In support of a critical realist perspective, and given the methodological steps emerging, constructive ‘trading zones’ between economic geography and aspects of heterodox economics may also be usefully pursued (Barnes and Sheppard, 2010).

Acknowledgment

The author would like to thank the reviewers and editor for their constructive engagement on the paper. Andy Pike, Robert Huggins, Duncan Maclellan, Pedro Marques, Tony O’Sullivan and Ken Gibb helpfully commented on earlier versions of the paper, and thanks also go to colleagues in Urban Studies - Simon Joss, Moira Munro, Rafael Verduzco Torres and Adriana Mihaela Soaita - for their constructive inputs at various points. Parts of the paper draw on doctoral research undertaken at the University of St Andrews. The paper reflects the views of the author only.

Declaration of conflicting interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Aspects of this paper draw on research undertaken as part of a wider doctoral project which received support from the University of St Andrews and the City of Edinburgh Council.

ORCID iD

David Waite https://orcid.org/0000-0003-0567-5683

Notes

1. Allana and Clark (2018: 1) give the following definition of meta-theory: ‘refers to broad perspectives, which make claims regarding the nature of reality. Meta-theories philosophically underpin research and practice’.
2. Such as their alleged ‘horizontalism’ (including the absence of social hierarchy) and their ‘black-and-white’ approach to characterising the urban.
3. This paper consciously leaves to one side more normative questions around the desirability of growth.
4. This is a broad range identified in the literature; though there are outliers.
5. The employee job data are 2020; the GVA data are 2019 (Business Register and Employment Survey (NOMIS) and Regional Gross Value Added - Sources: Office for National Statistics).

References


Centre for Cities (2022) City factsheet - Edinburgh. Centre for Cities. Available at: https://www.centreforcities.org/city/edinburgh/ [retrieved 18/05/22].


