Abstract

The article asks why institutional reforms work in one place and not another, and why old ways of doing things can prove so resilient. It argues in favour of a concept of institutional formation, which is different from ‘institutional design’ as a time-limited event or ‘institutional change’ as an open-ended historical trajectory. Institutional formation is conceptualised as an animated, nested and embedded process. A multi-level framework is developed that specifies the links between institutional actors, institutional rules and institutional contexts. The model is elaborated with reference to a case study of local government reform in England, specifically the devolution of responsibilities from central government to voluntary collaborations of elected local authorities (‘combined authorities’). The model is used to explain variation in the process of institutional formation in two different city-regions, focusing on the role of leaders, legacies and localities.
Introduction

Why do institutional reforms work in one place and not another? Why do some initiatives never make it off the drawing board or have a limited life? Why do old ways of doing things appear so resilient in the face of new designs? Such questions have great significance given the importance of institution building to tackle contemporary governance challenges, whether through reforming established structures or setting up new arrangements. To address these questions we need to bridge two established academic literatures - on institutional design and institutional change – which conventionally remain at a distance from one another. This article argues in favour of a concept of *institutional formation*, which is different from institutional design as a time-limited event or institutional change as an open-ended historical trajectory. Institutional formation is conceptualised as an animated, nested and embedded process. The article starts by developing a model of institutional formation which specifies the links between institutional actors, institutional rules (at different levels) and institutional contexts. The model is then elaborated with reference to a case study of local government reform in England, specifically the devolution of responsibilities from central government to voluntary collaborations of elected local authorities (‘combined authorities’). The model is used to explain variation in the process of institutional formation in two different city-regions (Greater Manchester and the West Midlands).

Conceptualising institutional formation

Our concept of institutional formation is grounded in a pragmatic epistemological position that enables the crossing of theoretical boundaries (rather than spawning more ‘schools’ of institutionalism). The ‘institutional design’ literature expresses a rational choice orientation in which actors are able to craft institutions to address collective action problems, through the manipulation of rules, information and incentives (Kiser and Ostrom 1982). With its voluntaristic stance, such an approach can be seen as under-estimating the limits to design and as displaying an excess of heroism. On the other hand, the ‘institutional change’ literature (associated with historical institutionalism) focuses on the power of path dependency to undermine attempts at intentional institutional change, either strangling them at birth or derailing them along the way (Pierson 2004). With its emphasis upon the constraining influence of existing structures, such an approach can be seen as over-estimating the limits to design and as displaying an excess of modesty. By bringing together insights from both literatures, our model provides a more nuanced understanding. In fact, the two literatures show signs of moving closer together, although this is not always acknowledged. Design approaches are reconsidering the role of contextual factors, through the effect of repeated interactions over time and the role of cultural norms (Ostrom 2005, 27). Historical institutionalists increasingly argue for ‘bringing the actor back in’, showing how actors’ interpretation of rules is itself a source of change, however gradual (Steinmo 2008, Mahoney and Thelen 2010). The two literatures tend to remain divided, however, by their level of analysis (micro/macro), timescale (long/short) and their methodologies (modelling/case studies).

All ‘new institutionalists’ see political institutions as ‘more than’ organisations. Institutions constitute the ‘rules of the game’, which shape the choices and behaviour of individual and organisational actors. But while historical institutionalists see institutions as entire policy systems, institutional design scholars focus in on specific ‘action arenas’ within which actors encounter institutional rules. A distinguishing feature of new institutionalist scholarship, from March and Olsen (1984) onwards,
has been its recognition of the role of informal elements in bolstering institutional regimes. And yet, historical institutionalists often argue that the informal remains outside their sphere of study (due in no small part to the difficulty of researching this in historical perspective) (Streeck and Thelen 2005, 10), while institutional design scholars concern themselves chiefly with that which is more open to human intervention (i.e. formal rules and structures) (Kiser and Ostrom 1982). Helmke and Levitsky (2006) have studied the contradictory and consolidatory effects of separate formal and informal institutions within entire political systems (e.g. clientelism and representative democracy). This article argues, however, that all institutional formations are characterised by a combination of formal rules (including structures and processes) and informal conventions (established and routinized practices): this is what Elinor Ostrom (1999) refers to as ‘rules-in-use’. We add to Ostrom’s formulation by arguing that such configurations of rules are backed up by customary narratives, which elaborate the underpinning ideas – the ‘reasons why’ institutions operate as they do (Lowndes and Roberts 2013; Schmidt 2008).

Rules, practices and narratives are thus three, interlocking, modes of institutional constraint through which the choices and behaviours of actors and organisations are shaped. Indeed, it is the combination of multiple means of constraint that makes institutions ‘special’, giving them their capacity to endure over time and also to generate loyalty and affection from actors, rather than relying solely upon force or incentive-based compliance mechanisms. Rejecting the invitation of the dominant institutionalist literature to ‘choose between’ causal mechanisms, this article sees institutions as shaping behaviour through the combination of regulative, normative and cognitive constraints, a mix that varies in distinct temporal and spatial contexts, and can only be established empirically (rather than predetermined ontologically). We also note that the sequencing of change in regulative, normative and discursive elements may be different. Modes of constraint may combine to produce institutional stability over time, but where gaps emerge (e.g. formal rules unsupported by practices or narratives), instability results. Such instability may lead to institutional collapse or, indeed, new opportunities for change (Lowndes and Roberts 2013, 69).

Institutional formation is a process through which new rules-in-use and supporting narratives are established. While institutions constrain human behaviour, they are themselves human constructions. This observation refers, of course, to grand foundational moments (writing constitutions, legislation or international agreements) and acts of major institutional reform. But it also draws attention to the day-to-day enactment of political institutions by agents. Rules-in-use are only ‘in-use’ because actors use them (this is the key difference from rules-in-form.) Actors reproduce rules as they use them, thus generating the regularities and stability associated with institutions (the ‘quasi parameters’ within which institutions can become self-reinforcing according to Greif and Laitin, 2004). But, at the same time, interpretation and enactment are sources of institutional change. Actors have to fit situations to rules and, in so doing, are likely to adapt those rules (particularly as contexts change), and they may break or subvert rules in pursuing their own interests. As Mahoney and Thelen (2010, 7-8) observe, institutions are ‘distributional instruments laden with power implications’. Historic compromises between groups may build in tensions that later produce pressures for institutional change; or new circumstances may increase the authority and capacity of certain groups of actors to direct change.

Institutional formation does not unfold due to unflinching logics of path dependency, whereby early institutional choices constrain subsequent decisions (the cost of changing path being higher than any
inefficiencies associated with remaining on track). Neither can it be explained by what Bob Goodin (1996, 28) calls the ‘Myth of the Intentional Designer’. Power and agency are exercised within institutional settings characterised by specific sets of linkages across two dimensions. First, any set of rules-in-use is vertically-nested ‘within an ever-ascending hierarchy of yet-more-fundamental, yet-more-authoritative rules and regimes and practices and procedures’ (Goodin and Klingeman 1996, 18). Kiser and Ostrom (1982) distinguish between operational (day-to-day), collective (legislative or policy based) and constitutional rules. The parameters of institutional formation at one level are constrained by higher level rules (although these too are subject to the effects of power and agency). Second, rules-in-use are horizontally-embedded within wider institutional contexts. The literature on ‘varieties of capitalism’, for instance, demonstrates how capitalism takes particular forms in countries with different social and political institutions (e.g. parties, trades unions, religious blocks) (Hall and Thelen 2008). Explaining processes of institutional formation requires that we understand not just path dependency but spatial contingency. Institutional choices in the past may influence a particular process of institutional formation, but so might contemporaneous linkages with ‘neighbouring’ institutions (e.g. between health care and taxation institutions at the national level, or between the institutions of municipal government and civil society). No institution stands alone, but is interconnected with a range of other institutions, which reinforce or undermine its effects; institutional configurations may reflect ambiguous accommodations as well as complementary relationships (Lowndes and Roberts 2013, 42, Mahoney and Thelen 2010, 22).

Both path dependency and spatial contingency can provide a brake on change, but they can also provide resources for actors engaged in processes of institutional change. The range and character of institutional materials ‘to hand’ are specific to each institutional context, as is the balance of power between actors seeking to mobilise them. We can identify three typical strategies: institutional remembering (where redundant, subordinate or latent institutional paths are reactivated for new purposes), borrowing (where rules, practices or narratives are transferred from adjacent institutional domains) and forgetting (where active maintenance or enactment of existing institutions is withdrawn) (Lowndes 2005, Lowndes and Roberts 2013). This typology recognises that there is rarely just one institutional path (Crouch 2005). More likely, there are multiple paths which are more or less entrenched at any one time, and whose significance varies over time depending on the level of active support they receive from dominant actors and their resonance (or otherwise) with changing aspects of the institutional configuration (e.g. higher level constitutional or legislative rules, or changes in associated economic or civic institutions).

In summary, our concept of institutional formation is different from design as a time-limited event, but also from institutional change as an open-ended historical trajectory. Building on areas of overlap between rational choice and historical institutionalism, the purpose of the concept is to frame research on intentional processes of institutional development in such a way that serious attention can also be paid to temporal effects (without engaging in a chiefly historical endeavour). From rational choice institutionalism, we take specifically the concept of rules-in-use, which are animated by individual actors, composed of formal and informal elements, and vertically nested in multi-level institutional frameworks. Historical institutionalism, on the other hand, informs our understanding of the horizontal embeddedness of these rules (and supporting narratives), and the way in which both path dependence and spatial contingency influence institutional change, in a context of shifting power relations. As shown in Figure 1, the process of institutional formation is
influenced by actors themselves, higher level institutional rules and wider temporal and spatial contexts. The model is underpinned by three propositions:

- Institutional formation is a *nested* phenomenon, influenced by multi-level relationships between constitutional, governmental and operational rules
- Institutional formation is an *embedded* phenomenon, influenced by temporal and spatial contextual effects
- Institutional formation is an *animated* phenomenon, influenced by actors’ interventions, interpretations and contestations

Next we explain our research design and introduce the case study, which will enable the testing and further development of the model.

*Figure 1 about here*

**Research design: Case study of city-region devolution**

As a research arena for the study of institutional formation, local government has the benefits of propinquity and numerosity (John 2009), providing many accessible and comparable units for analysis, which show institutional diversity within a single legal and constitutional framework. Local government is also recognised for its capacity to generate institutional innovation, given its relative distance from the imperatives of the nation state and its closeness to distinctive local communities (Dryzek 1996). Elinor Ostrom (2010) explains in her Nobel Prize lecture how her work on institutional design started with comparative studies of US metropolitan governance. English local government has pioneered many institutional reforms, including new public management, public/private partnerships and co-production of services with communities. From 2010, however, the agenda has been dominated by central government austerity policies, with local authorities losing 30% of their revenue budgets under the Coalition government, and the 2015 Conservative government pledging a further cut of 56% (HM Treasury 2015a). Our case study concerns the devolution policy introduced against this backdrop, championed between 2010 and 2016 by Chancellor of the Exchequer, George Osborne.

Through a series of negotiated ‘settlements’, a range of powers and resources have been devolved from central government to voluntary collaborations of elected local authorities (‘combined authorities’) working at a city-region level. City-regions cover a major city and the neighbouring towns and suburbs that constitute its ‘functional economic geography’ (Ward 2010). Devolution was aimed at stimulating economic development outside London and the South-East through the creation of ‘growth hubs’. George Osborne (2014) argued for the creation of a ‘Northern Powerhouse’ of city-regions and, subsequently, a similar ‘Midlands Engine’. Unlike previous attempts at introducing regional governance, each devolution settlement involved the creation of a bespoke package of powers. Without an elected assembly, combined authorities are governed by a cabinet made up of the political leaders of the constituent bodies; for the most ambitious devolution package, central government also required the new office of a directly elected mayor (Gains 2015). After a series of individual acts of parliament to establish the first combined authorities, enabling legislation was passed (Cities and Local Government Devolution Bill 2016). Devolution was widely seen by local authorities themselves as the only route to accessing additional resources in the context of austerity.
We compare the process of institutional formation in two different city-regions, Greater Manchester and the West Midlands, which were selected as ‘most different’ sub-units within an ‘embedded case study’ of the formation of combined authorities (Yin 2014). The two city-regions are similar in terms of size and economic significance; indeed, Manchester and Birmingham (the major cities within each region) vie with one another for the status of England’s ‘second city’. Yet the experience of forming combined authorities has been very different. Greater Manchester is the furthest advanced example of devolution, while the West Midlands has experienced a more faltering and contested process. The case study follows an ‘explanation building’ logic whereby initial theoretical propositions are compared with findings from the case and revised as appropriate (for application to future cases) (Yin 2014). The aim is to ‘develop case-specific observable implications’ of the theory in question (Bennett and Checkel 2015, 18). The case study provides an opportunity to operationalise and elaborate the theoretical propositions, and to assess the extent to which they help explain variation in the process of institutional formation. While the research design does not enable the testing of a series of rival hypotheses, it does address a gap in the existing literature, which focuses on evaluating the extent of power devolved by central government (seeking to compare statements of intent with emerging outcomes) and assessing the relative significance of its economic and political strategies. Explanations of the devolution phenomenon pay little attention to local level variables, concentrating upon central government strategy, specifically the enduring centralism of the ‘British Political Tradition’ (Richards and Smith 2015) and the historic weakness of elected local government (Lee 2017, Hambleton 2017, Blankett et al 2016). There is a lack of systematic research comparing the process of institutional formation in different city-regions or seeking to explain why local outcomes vary.

The research employs a process tracing methodology, defined by Collier (2011, 823) as ‘the systematic examination of diagnostic evidence selected and analysed in the light of research questions’ posed by the investigator. Collier (2011, 824) also notes that ‘prior knowledge’ is vital for effective process tracing - not just about the conceptual framework but also the empirical regularities within the case that enable the researcher to establish starting points and select critical episodes and events. Our prior knowledge was developed through participant observation at four workshops convened by our University for politicians, public servants and community activists involved in setting up combined authorities (approximately 100 participants). Through wide ranging conversations regarding the challenges and opportunities of the new governance arrangements (conducted under Chatham House rules), we were able to start mapping the distinctive ‘processes, sequences, and conjunctions of events’ (Bennett and Checkel 2015, 7) in our two city-regions. This prior knowledge was then used to identify ‘diagnostic evidence’ upon which to base inferences about the sources of variation within the process of institutional formation. A total of 95 documents were analysed, including combined authority documents, legislation and government reports, speeches from local and national politicians and public servants, think tank papers, and local and national media (print and online). The data were analysed using NVIVO, with primary codes based upon the three theoretical propositions. Data were collected for the period up to June 2016, which marked the formal establishment of the second of the two combined authorities. We do not consider the subsequent operational dynamics of the two combined authorities, nor the formation of the other five combined authorities (see House of Commons 2015, 17-18).

We now present our research findings for Greater Manchester and the West Midlands, and then move on to analyse the contrasting processes of institutional formation using our conceptual model.
The Greater Manchester Combined Authority: ‘We want our city back’

The Greater Manchester Combined Authority was formed in 2011 and includes ten constituent local authority members (Manchester, Salford, Trafford, Bury, Stockport, Rochdale, Oldham, Bolton, Tameside and Wigan). The institutional antecedents in Greater Manchester supported the formation of a combined authority. With the abolition of Greater Manchester County Council in 1986, joint services (transport, fire, police and waste) were reorganised around a number of joint boards, overseen by a coordinating committee. Whereas in many places across the country this joint working was beset by political rivalry, politicians from across Greater Manchester showed willingness to cooperate, institutionalising the committee as the Association of Greater Manchester Authorities (AGMA) (Kenealy 2016). The largest and most important entity, Manchester City Council, was incentivised to collaborate due to its own institutional peculiarity as an ‘underbounded’ local authority area (leaving many of its suburbs outside the city boundary. While the formal institutional architecture was minimal, the arrangement maintained and fostered norms of trust and practices of collaboration among constituent authorities, whilst expressing a shared Greater Manchester identity (Hebbert and Deas 2000, 87). This distinctive institutional framing extended beyond AGMA itself, shaping the behaviour of actors in both local government and the private sector over a thirty year period. Collaborative initiatives included the Manchester Investment and Development Agency Service, the expansion of Manchester Airport, and hosting the Commonwealth Games (Hebbert and Deas 2000, 90). The city-region became increasingly interconnected along many of the economic dimensions that would subsequently underpin the combined authority.

In 2009, AGMA initiated an institutional review (led by its own think tank, New Economy) to explore the case for further collaboration, recommending a formalisation of arrangements and the establishment of a general power of competence (Manchester City Council 2010). At the same time, a joint elite of political and business leaders (the infamous ‘Manchester Men’) lobbied central government; they persistently argued the case for devolution of powers and funding (Kenealy 2016), winning the country’s first City Region Pilot (2009). Effectively an institutional prototype for the combined authority, the Pilot enabled the ten boroughs to pool responsibilities over housing, skills training, economic development and transport in return from devolved funding and powers. The negotiations directly influenced the 2009 Local Democracy, Economic Development and Construction Act, which provided the legislative underpinnings for the first combined authorities.

Under the Coalition government, the Greater Manchester Combined Authority Order passed through parliament in 2011. The combined authority reflected a continuity of membership with the old Greater Manchester County Council and sustained through AGMA. There was a virtually seamless transition from a voluntary federation of ten local authorities to a formal integrated governance arrangement. Institutional arrangements developed from the ‘bottom up’ rather than being imposed by Whitehall, having evolved over time to meet the needs of the city-region (whilst also shaping the national debate about devolution). With its combined authority up and running, Greater Manchester was able to integrate new national initiatives and opportunities into this institutional framework, including a Local Economic Partnership (LEP), of local businesses, required from 2011. In other city-regions, protracted negotiations between potential members, within a fragmented institutional landscape, had the effect of delaying and complicating the process of forming a combined authority.
Institutionalised collaboration in Greater Manchester was maintained after the end of the Metropolitan County Council not just through the enabling framework of formal rules provided by AGMA, but also through normative expectations and a powerful narrative around the city’s identity and potential – in its own right and as an alternative economic and political hub from London and the South East. These provided a ‘soft’ institutional legacy upon which the new formal architecture of the combined authority could be constructed. The narrative, or long-standing ‘Manchester script’ (Quilley 1999, 185), emphasised ‘urban entrepreneurialism’, seeking to build upon the locality’s industrial legacy whilst accepting the reality of post-war deindustrialisation (as embodied in the Manchester Independent Economic Review) (McKillop et.al. 2009). A clear vision was expressed: the combined authority was a vehicle to ‘take back control’. The years without strong city-region governance were cast as an exception, rather than the norm. The challenge was to rearticulate the Manchester script in the context of a new economy. Greater Manchester’s history was that of a series of connected conurbations (serving the textile trade in particular) that ‘became the hub of the world’s first industrial metropolis’ (Hebbert and Deas 2000, 81). Sir Howard Bernstein and Sir Richard Leese, the chief executive and political leader of Manchester City Council, argued it was ‘absurd that a city that had once built and run its own schools, hospitals, museums, transportation and social services should languish under the lash of Whitehall’, with Bernstein adding ‘we want our city back’ (Jenkins 2015). In this way, the Manchester script contributed to the construction and narration of Chancellor George Osborne’s higher level institutional discourse of building a ‘Northern Powerhouse’ of city-regions (which even made it into the 2015 Queen’s Speech).

But the availability of an institutional narrative did not determine the formation of the GMCA. The narrative was creatively re-worked and persuasively articulated in order to open the ‘window of opportunity’ presented by the Chancellor’s devolution initiative, whilst also gaining traction over potentially competing narratives which highlighted poverty and inequality (rather than economic growth) within the region (Coleman et al 2015). Devolution was neither a uniform policy nor an inevitable outcome for all areas across England; rather, it was an outline institutional framework for bilateral negotiations between specific city-region leaders and Whitehall. Greater Manchester’s achievements can only be explained by looking at the relationship between longstanding (and evolving) institutional frameworks and the sustained strategic agency of key actors. The case for the Greater Manchester Combined Authority was developed over many years and not formulated swiftly or simply in response to calls from central government for devolution (GMCA 2013). AGMA had created New Economy (and seven ‘policy commissions’) to explore and articulate strategies for economic growth, public service reform and city-region governance. The Treasury was prepared to work closely with Greater Manchester, which became a test-bed for new ideas about devolution within a multi-level and dynamic process of institutional formation. Further powers were subsequently devolved, including aspects of health, housing and policing, via four subsequent devolution deals (November 2014, July 2015, November 2015 and March 2016). Government willingness to accommodate Greater Manchester’s ambitions was evident when the Cities and Local Government Devolution Bill (2016) made it legally possible to devolve to a combined authority any function of a constituent council or public authority.

Negotiations relied upon a high level of personal trust between civil servants in the Treasury and the leadership in Greater Manchester. Bernstein and Leese acted, in effect, as institutional entrepreneurs, becoming joint authors with central government of the combined authority/devolution model. Meetings with the Treasury were frequent, but conducted largely in
secret, and without proposals being subject to consultation within the wider Greater Manchester population (Richards and Smith 2015). While it has been claimed that Bernstein ‘got everything he asked for’ (Jenkins 2015), there was also compromise. The government made devolution conditional upon a directly-elected mayor, who was to act as a single point of accountability (Osborne 2014). This had been a major sticking point amongst local leaders, not least because Manchester itself had rejected the idea in a 2012 city referendum. Ultimately agreement was reached that a mayor would be accountable to a cabinet of leaders from the ten constituent authorities (HM Treasury 2014), going some way to protecting the power base of the ‘Manchester Men’.

The West Midlands Combined Authority: ‘Playing catch-up’

The West Midlands Combined Authority (WMCA) was established in 2015 and includes seven constituent local authority members (Birmingham, Solihull, Coventry, Wolverhampton, Sandwell, Dudley and Walsall) (HM Treasury 2015b). In contrast to the GMCA, it also includes non-constituent authorities with limited voting rights. Prior to the formation of the West Midlands Combined Authority, the region was characterised by complex institutional arrangements with overlapping memberships and responsibilities (Kerslake 2014, 7). With the abolition of the West Midlands County Council in 1986, joint-decision making responsibilities were assigned to a number of joint committees, with varying memberships, including the Black Country Joint Committee, the West Midlands Joint Committee, an Integrated Transport Authority and joint boards for shared emergency services. This institutional arrangement both reflected and propagated historical inter-regional tensions and rivalries, most acutely between the Black Country authorities (Sandwell, Wolverhampton, Walsall and Dudley) and Birmingham City Council, reflecting different industrial and social histories dating from the Industrial Revolution (Spencer 1986).

It was not until 2014 that local authorities in the West Midlands began to consider seriously the benefits of greater integration, lacking the institutional continuity provided by AGMA in Greater Manchester. The formation of a combined authority offered the only route for accessing new funding from central government in the face of deepening austerity, and there was an increasing feeling that the region was getting ‘left behind’ others in the North of England (Bailey 2014). But the timing of institutional changes proved significant. For Greater Manchester, the early adoption of the combined authority model meant that it could assimilate the new institutional requirement for a Local Enterprise Partnership. In the West Midlands, the LEP initiative pre-dated the formation of the combined authority, and its implementation reflected, and reinforced, dominant institutional rules, practices and narratives that that were not conducive to city-region integration. Instead of one West Midlands LEP, three were established, expressing in a new institutional form old rivalries, identities and industrial traditions. Each LEP produced its own strategic economic development plan, leading to three separate City Deals with central government. When the WMCA was eventually established, it also included a fourth LEP, bringing its total local authority membership (constituent and non-constituent) to twenty-two. In the West Midlands, institutional legacies (formal and informal) provided an obstacle rather than a facilitator to the formation of a combined authority.

In both Greater Manchester and the West Midlands, markets for labour, goods, services and investment were no longer based upon small town geography; but it was Greater Manchester that had the institutional capacity to reflect and support these developments. Institutional arrangements in the West Midlands reflected anachronistic economic geographies, exercising path dependent
effects in favour of localised and competitive governance bodies. Hence the West Midlands was late coming to the negotiating table and experienced difficulties articulating a case for devolved power, given the formal separation of the different LEPs (and their City Deals) and the lack of a persuasive shared narrative. The process of institutional formation was characterised by conflict as well as fragmentation. The Black Country local authorities and Birmingham City Council signed an agreement in November 2014 to develop a combined authority, and started talks about the possible extension of membership to neighbouring towns, Solihull and Coventry (seeking to deploy the institutional memory of the old West Midlands Metropolitan County Council) (Dale 2014a). Considerable opposition was launched by Solihull, which was distrustful of Birmingham City Council’s ability to work collaboratively (Birmingham Post 2015), given a highly critical central government review (Kerslake 2014). Coventry, while less hostile, was unable to join the combined authority without Solihull because the 2016 legislation required that constituent bodies shared borders (Brown 2014a, Smulian 2014); residents’ sense of attachment to the neighbouring rural county added a further obstacle.

In contrast to the leading role played by Manchester City Council in the combined authority negotiations, the process in the West Midlands was only unblocked when Birmingham City Council agreed to stand back (acknowledging criticisms of its institutional health and collaborative capacity). While Birmingham’s Chief Executive continued to play an important role ‘behind the scenes’, the higher profile positions were taken by smaller partners. Birmingham City Council’s situation was shaped by prior institutional considerations, notably the fact that (in contrast to Manchester City Council) it is a notoriously over-bounded local authority, covering a population of approximately one million and absorbing many of the city’s suburbs (sometimes against popular sentiment). The combined authority negotiations were in part the outworking of this prior institutional choice. Institutional rules, practices and narratives reinforced a Birmingham-only approach, and its sheer size evoked hostility from neighbouring local authorities that feared Birmingham’s dominance of any new combined authority (Birmingham Post 2014). Equivalent fears were not entirely absent in Greater Manchester, but were accommodated within a longer term process of institutional formation. Unlike Birmingham, Manchester City Council was not tainted by proven institutional weakness and, as the largest actor, was able act as a facilitator of collaboration (rather than an impediment), both within the city-region and with central government. The lack of a pre-existing West Midlands identity made difficult the articulation of an institutional narrative to support the new combined authority. Consequently there was no clear framing of questions about membership, boundaries and objectives (reflected in the fierce debate over a name for the combined authority). The previous West Midlands Metropolitan County Council had passed into history without leaving the strong institutional legacies that were exploited in Greater Manchester. Instead, actors faced the challenge of reinventing what was widely regarded as a failed institutional identity.

The eventual proposals for a West Midlands Combined Authority followed closely central government preferences. The complex and contested negotiations within the city-region had delayed engagement with central government, during which time the national negotiating environment had significantly changed. Both timing and sequencing were important: the West Midlands’ experience was different because it came after Greater Manchester. As the National Audit Office pointed out, localities increasingly ‘looked to precedent in the form of deals already agreed in other areas’ (NAO 2016, 9). At the same time, central government had ‘strengthened its management approach’ (NAO 2016, 10), appearing to lose interest in principles of bespokeness and
reverting to more established institutional practices of central direction and uniform local implementation. The West Midlands bid was prepared in just six weeks, following the July 2015 Spending Review that invited a new tranche of proposals by the start of September. The extremely limited time available for discussions with the Treasury contrasted with the longer term, organic and bottom-up developments in Greater Manchester. The historic institutional fragmentation that characterised the West Midlands had also precluded the emergence of a strong leader (or joint elite as in Greater Manchester) to advocate for devolution and build relationships with national politicians and civil servants. The WMCA vision appeared motivated in a large part by the need to ‘catch up’ with other city-regions (notably Greater Manchester); neither was it supported by a powerful narrative at the national level, with George Osborne’s more modest ‘Midlands Engine’ failing to gain the traction of the earlier Northern Powerhouse discourse.

**Explaining variation in institutional formation**

Using a process tracing methodology, we have compared the experience of institutional formation in two city-regions. We have seen how Greater Manchester pioneered the combined authority model (while the West Midlands played a reluctant game of ‘catch up’) and negotiated a series of devolution deals with central government. Devolution in Greater Manchester involved aspects of health and social care, police and fire services, and criminal justice, in addition to the functions that characterised the West Midlands deal (transport, housing, further education, employment and business support) (NAO 2016, 4304). Despite what Harrison (2015) calls ‘headline similarities’, the process of institutional formation in Greater Manchester had made considerable inroads into core public services. The combined authority became more than an economic development vehicle prescribed by central government.

We return to our three propositions in order to explain how and why institutional formation differed in the two city-regions. A revised model (Figure 2) shows the process of institutional formation for our case study.

*Figure 2 about here*

(a) Institutional formation is a nested phenomenon, influenced by multi-level relationships between constitutional, governmental and operational rules

The process of establishing combined authorities was influenced by multi-level relationships between constitutional, governmental and operational rules. The devolution policy was, at first sight, surprising given entrenched institutional centralisation (with low levels of local government autonomy), the abolition of metropolitan government in the 1980s, and the subsequent (and repeated) failure to embed new forms of English regional governance. Our model helps explain this anomaly. To understand the emergence of new combined authorities at the operational level, we need to look at how established governmental rules were themselves destabilised. We need to look both ‘upwards’ to perturbations at the constitutional level, and ‘sideways’ to changing institutional configurations in the economic arena.

One of the effects of the closely fought (though unsuccessful) referendum for Scottish independence (September 2014) was to challenge the central/local settlement within England. With the increase of devolved powers to Scotland as part of the post-referendum settlement, demand was fuelled
among the leaders of major cities in the North of England for some degree of English devolution (Core Cities 2013, Lowndes and Gardner 2016, 361-2). The Chancellor of the Exchequer, George Osborne, took control of the English devolution agenda, seeking to create not just new formal institutional vehicles but supportive institutional narratives (the Northern Powerhouse and, less successfully, the Midlands Engine) (Osborne, 2014). Devolution was also shaped by the government’s own austerity agenda, providing as it did an opportunity to shift downwards responsibility for implementing spending cuts (which far outstripped the new resources on offer to combined authorities) (Lowndes and Gardner 2016, 365).

These new top-down impulses for institutional reform met longstanding bottom-up campaigns for greater local government autonomy. Not only were political incentives (and opportunities) changing, so too were economic imperatives. In the context of sluggish economic growth and a commitment to pay back the national deficit, the Chancellor’s growth strategy required the alignment of governance and economic institutions at the regional level, with the aim of promoting growth hubs beyond London and the South East (City Growth Commission 2014). City-region governance, which could oversee infrastructure, training and business support functions, was seen as necessary to catalyse and capture ‘benefits of agglomeration’. As Figure 2 shows, the process of institutional formation was not only vertically nested, but horizontally embedded too.

While coming after the establishment of combined authorities in Greater Manchester and the West Midlands, the referendum decision for Britain to leave the EU in June 2016 destabilised constitutional rules further. It remains to be seen whether the outcome will be a strengthening of institutional paths in favour of a strong unitary state, or the amplification of revolutionary paths as pro-EU Scotland pushes for greater autonomy from Westminster. As our research shows, there is never just one institutional path, but rather a range of potential institutional frames. Subordinate paths are capable of strategic rehabilitation, while dominant paths may be vulnerable as contexts and power relations change. In keeping with our model, changes at the constitutional level can be expected to reverberate through all institutional levels – both political and non-political (including the economy and civil society). The post-referendum government of Theresa May was less enthusiastic about the further spread of city devolution, and the policy lost a critical ‘institutional entrepreneur’ with the sacking of Chancellor, George Osborne. The government’s austerity programme, which had shaped both the design of devolution and local authorities’ response, was also institutionally weakened with the post-Brexit abandonment of previous deficit reduction targets and the subsequent failure of the Conservatives to gain a parliamentary majority in the 2017 general election. Future research will establish whether the new operational rules underpinning combined authorities will prove sufficiently robust to withstand new developments at the constitutional and governmental levels, and indeed how they might continue to change over time.

(b) Institutional formation is an animated phenomenon, influenced by actors’ interventions, interpretations and contestations

The role of local leaders was a critical factor in the institutional formation process in Greater Manchester and the West Midlands; indeed, locally embedded processes of elite formation and strategic action help explain different outcomes in the two city-regions. In neither instance was there any wider or sustained engagement with non-elite actors, including members of the public and community groups (NEF 2016, Blunkett et.al. 2016). Even backbench councillors in the constituent
local authorities that came together to form the combined authority were marginalised (Dale 2015). In Greater Manchester, leadership was more clearly defined, with obvious spokespeople and negotiators and a high degree of integration between political and economic elites at the local level (consolidated over time) and between local elites and national politicians and civil servants (House of Commons 2015, Jenkins 2015). Leaders’ capacity was underpinned by formal institutions (AGMA and Manchester City Council as the lead player), informal conventions of collaboration, and a powerful shared institutional narrative (‘we want our city back’) (GMCA 2014).

Lobbying of central government took place within a relationship of trust and reciprocity; it was not just reactive (to get the best deal) but also proactive. Leaders worked with central government to shape not just operational rules for Greater Manchester but also the governmental rules themselves (i.e. the devolution policy, the narrative of the ‘Northern Powerhouse, and the associated legislation) (House of Commons 2015). The strategic agency of Greater Manchester’s institutional entrepreneurs had knock-on effects at the constitutional level, through destabilising established practices of central/local relations and offering a response to the ‘English question’ (less challenging to national politicians than alternative proposals like an English parliament). Reflecting these research findings, the arrows linking institutional levels in Figure 2 (above) have been made double-headed, showing that institutional nesting can be a two-way process of mutual influence rather than the one-way typically envisaged in theories of institutional design, which sees higher level rules as necessarily ‘more authoritative’ (Goodin and Klingemann 1996, 18).

Leadership was far more contested in the West Midlands, reflecting the fragile character of formal institutional arrangements for joint-working, and enduring informal practices of competition and distrust (Kerslake 2014, Game 2014). The lack of an obvious role for the major player in the city-region, Birmingham City Council (due to its own institutional failures and a lack of support from neighbouring local authorities) further undermined leadership capacity, as did the lack of publicly identifiable spokespeople and negotiators (Blackett 2014) and the weaker integration of political and economic elites (Brown 2014b). The lack of a history and culture of collaboration and of a shared institutional narrative (beyond ‘playing catch up’) was reflected in conflict over the membership, boundaries and purpose of a new combined authority, showing the weakness of processes of institutional formation that rely on formal rule change alone (Bailey 2014, Brown 2014a). The lack of clear and consolidated leadership left the West Midlands vulnerable to being rushed by central government into adopting the default option of the (initial) ‘Manchester model’, which had become over time an institutional template (despite the founding claims of the devolution policy) (Dale 2014b). In conditions of uncertainty, as DiMaggio and Powell (1991, 64) argue, organisations tend to borrow institutional rules from those bodies that appear to be coping most effectively with change (via ‘isomorphism’) - whether or not they are best suited to their own specific needs.

Reflecting on the challenge of brokering deals, Harrison (2015) notes that there was ‘inevitably a limit to the amount of times this can be achieved… with different places, within a given period of time’. While this is hard to represent in Figure 2, we have made the arrow linking leaders and institutional formation doubled-headed in order to demonstrate that the devolution process itself influenced leadership dynamics within localities, as well as vice versa. The obligation to introduce a directly elected mayor in 2017, as part of the devolution deals, has influenced these dynamics further still.
(c) Institutional formation is an embedded phenomenon, influenced by temporal and spatial contextual effects

The role of context in shaping the process of institutional formation, identified in Figure 1, was strongly confirmed in our research on Greater Manchester and the West Midlands. Reflecting the research findings, Figure 2 separates out the influence of spatial and temporal effects. While analytically separate, localities and legacies are, in practice, deeply entwined; hence the double-headed arrow linking the two variables in Figure 2. The arrows linking localities and legacies to institutional formation are also double-headed, so as to reflect the impact of the new combined authorities on these factors, as well as vice versa.

In Greater Manchester, institutional rules associated with the former Metropolitan County Council had been maintained through the formal architecture of AGMA, and associated informal practices that became embedded over time and were enacted via a range of collaborative initiatives (GMCA 2014, Hebbert and Deas 2000). These institutional resources, along with the ‘Manchester script’, formed the underpinnings of the new combined authority. Rules and practices favouring collaboration were recuperated and rearticulated in the service of the campaign for city-region devolution (Blond and Morrin, 2014.). In the West Midlands, the former Metropolitan County Council had been unceremoniously buried, with few institutional resources remaining. Overlapping and uncoordinated joint boards provided the only formal institutional residue, while informal practices and narratives reflected a distrust of collaboration, serving instead to reinforce multiple identities (Walker 2014). Path dependencies and spatial contingencies provided resources for institutional formation in Greater Manchester (the under-bounded nature of the city council, the formation of AGMA, and a common industrial heritage); in the West Midlands, however, they provided obstacles (the over-bounded nature of the city council, fragmented institutional arrangements for joint working, competing economic geographies, and discrete industrial and cultural heritages) (Game 2014). Our analysis has shown how the different institutional legacies of the two city-regions shaped the process of forming a combined authority. But it also demonstrates that actors can work strategically with institutional resources, which shape rather than determine outcomes. In Greater Manchester, leaders were engaged in a process of active institutional remembering, whereby a metropolitan institutional path that had become subordinate over 30 years was rehabilitated as a city-region framework, enabling the adaptation to circumstances and maximisation of new opportunities. Locality effects are important in explaining the different processes of institutional formation in the two areas; however, these refer not to geography per se but to the different institutional framings of political and economic space over time.

In the West Midlands, the metropolitan path had, in contrast, been subject to processes of institutional forgetting whereby the active maintenance of collaborative institutions was neglected after the abolition of the county council. Neither were rules-in-use favouring collaboration cultivated via institutional borrowing from the economic sphere (as in Greater Manchester), where markets were steadily expanding beyond the ambit of individual local authorities. Using Mahoney and Thelen’s (2010) terminology, we can observe a process of ‘institutional conversion’ in Greater Manchester (new purposes attached to old institutions), in contrast to a delayed and then hurried attempt in the West Midlands at ‘institutional layering’ (new elements attached to existing institutions). The compromises and ambiguities built into the West Midlands model could, however, prove to be a source of dynamism in future rounds of institutional formation. Our research showed
that the less-than-wholehearted adoption of one particular institutional geography, and the insistent remembering of more localised community identities, political traditions and economic boundaries, served to constrain the process of forming a combined authority in the West Midlands. Yet these same institutional elements could provide valuable institutional resources in responding to new challenges as governance contexts change. Such institutional ambiguities reflect active processes of contestation between interests and identities within the city-region. It is important to note that the Greater Manchester ‘script’ is also not entirely uncontested: ‘The sceptics refer to the “northern poorhouse”’, highlighting poverty and health inequalities as well as the unequal burden of austerity measures on northern local government’ (Coleman 2015, 381).

Conclusion

This article has proposed a model of institutional formation that seeks to bridge the established academic literatures on institutional design and institutional change. Institutional formation is different from institutional design as a time-limited event and from institutional change as an open-ended historical trajectory. Institutional formation is conceptualised as a process through which new rules-in-use and supporting narratives are established, but within distinct institutional contexts and animated by particular sets of institutional actors (reflecting underlying power relations). The paper has tested the model through a case study of local government reform in England, specifically the devolution of responsibilities from central government to combined authorities.

The model demonstrates how operational rules emerging from a process of institutional formation are influenced by governmental and constitutional rules. It also shows how these operational rules can shape higher level rules through iterative relationships that develop over time, thus challenging simplistic concepts of institutional nesting. Comparing institutional formation in two different areas, we used the model to elucidate the factors accounting for variation. The agency of political leaders at the local level was found to be critical; citizens and communities were effectively shut out from what became a largely technocratic process. Where leadership was clear and sustained over time, actors were most able to influence the process, especially where there were consolidated elite relationships with local economic leaders and with national civil servants and politicians. The process of institutional formation rested upon the ways in which actors worked with existing rules-in-use in the process of forming a new combined authority. Where institutional legacies could be effectively mobilised in support of the new project (rather than presenting obstacles), the process was more robust and also more creative - in terms of innovating within (and influencing) the government’s institutional outline of what a combined authority might look like. While both city-regions had been subject to the same legislation in the past (notably the abolition of the metropolitan county councils in 1986), the nature and availability of institutional legacies depended upon the specific history, culture and identities of the locality in question. The formation of the new combined authorities was horizontally-embedded as well as vertically-nested, shaped by the opportunities (or obstacles) presented by unique local configurations of political and non-political institutions.

Underpinning our theoretical and empirical contribution is a determination to take seriously the relationship between institutional actors, institutional rules and institutional contexts. Our model provides a vehicle through which to analyse and compare processes of institutional formation in other settings – at the local, national or transnational level, in other policy arenas, and in other
country settings. It offers the possibility of identifying, and addressing, those factors and dynamics that account for the often unpredictable outcome of attempts at institutional reform. Our model is, of course, a simplification. In the light of our research, we have revised the original model to better reflect the elements and relationships at play. No doubt future research and real-world developments will lead to further iterations.
References


Birmingham Post. 2014. We’ll never be run from Birmingham - Coventry, 11 Dec.


Blackett, J. 2014. ‘Time to get on board, Coventry, or the region could be left behind’ Birmingham Post, 27th Nov.

Blond, P. and Morrin, M. 2014. ‘Devo Max – Devo Manc: Place Based Public Services’ Respublica


Brown, G. 2014b. Combined Authority’s a stich-up. Birmingham Evening Mail. 8 Nov.


Dale, P. 2014b. Why Birmingham Can’t Afford To Be Left Behind in the Metro Mayor Race’ The
Dale, P. 2015. ‘Devolution for Greater Birmingham ‘No Foregone Conclusion’, Warns Think Tank Chief. The Chamberlain Files, 28th Jan


Game, C. 2014. ‘If Manchester is the Mourinho of how to run things then we're in a lower league altogether’ The Birmingham Post, 28th Nov.

GMCA. 2013. Stronger Together: Greater Manchester Strategy 2013 (Manchester: Greater Manchester Combined Authority)

GMCA. 2014. A Plan for Growth and Reform in Greater Manchester (Manchester: Greater Manchester Combined Authority)


HM Treasury. 2014. *Greater Manchester Agreement: Devolution to the GMCA & Transition To A Directly Elected Mayor*. London: HMSO.


Methodologies in the Social Sciences, Cambridge: CUP


Walker, J. 2014 Nobody wants to destroy the identity of our area...we’re just saying together we are more than the sum of our parts’ *Birmingham Post*. 11 Dec.


Figure 1 – Model of the process of institutional formation
Figure 2 – Process of institutional formation for city-region devolution

Leaders
Clearly defined or contested political leadership for the city-region
Degree of integration of city-region political and business elites
Degree of trust in relationships with central government
Proactive or reactive lobbying on devolution policy
Strategic leaders negotiate deal, or deal imposed
City-region leadership underpinned by existing local institutions or not

Localities
Boundaries of existing political institutions impede or facilitate city-region working
Culture of collaboration, or distrust of joint working across city-region
Coordinated or fragmented approach to city-region governance
Congruence or dissonance of political and economic geographies in city-region
Clear and resonant place-based narrative or multiple competitive identities

Legacies
Institutional remembering or institutional forgetting
Redeployment of legacies of metropolitan governance or difficulty in rehabilitating
Level of active maintenance/adaptation over time of formal and informal collaborative legacies
Legacies facilitative or obstructive of devolution deal
Extent of institutional borrowing from businesses already working at city-region level
Mobilisation of specific place-based legacies within devolution narrative or adoption of template