

Urban Studies

<http://usj.sagepub.com/>

Why Does a City Grow? Specialisation, Human Capital or Institutions?

Michael Storper

Urban Stud 2010 47: 2027 originally published online 18 May 2010

DOI: 10.1177/0042098009359957

The online version of this article can be found at:

<http://usj.sagepub.com/content/47/10/2027>

Published by:



<http://www.sagepublications.com>

On behalf of:

Urban Studies Journal Limited

Additional services and information for *Urban Studies* can be found at:

Email Alerts: <http://usj.sagepub.com/cgi/alerts>

Subscriptions: <http://usj.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations: <http://usj.sagepub.com/content/47/10/2027.refs.html>

>> [Version of Record](#) - Aug 10, 2010

[Proof](#) - May 18, 2010

[What is This?](#)

Why Does a City Grow? Specialisation, Human Capital or Institutions?

Michael Storper

[Paper first received, July 2008; in final form, June 2009]

Abstract

Why are there persistent differences in income between metropolitan areas? The answer to this question has evaded much of the scholarship on the topic. Some of the frameworks that drive empirical research in this field are based on *ad hoc* combinations of explanatory factors, ranging from natural climate, to business climate, to land and labour costs. Theoretical approaches emphasise economic specialisation: some activities have higher rates of growth than others and this translates into divergence in interurban growth and income. Yet specialisation itself needs to be explained. International economics explains different growth rates and income levels among countries by emphasising specialisation, human capital and institutions. This framework can be adapted to the analysis of metropolitan growth. The thorniest aspect of doing so is to consider recursive relationships among the three, as well as decisive events that might introduce irreversible path-dependent outcomes that differentiate cities.

1. Introduction

Why does a city grow? Despite a long and venerable tradition, research on the growth of metropolitan area economies does not offer a convincing response to this question. Growth levels and per capita incomes exhibit strong and persistent differences across metropolitan areas. Within the US, for example, the San Francisco metropolitan area currently has a per capita income three times that of

Brownsville, Texas. These are obviously two metropolitan areas in structurally different development ‘clubs’,¹ but San Francisco also has a per capita income fully one-third greater than its southern neighbour, Los Angeles, which belongs to the same club.

It is surprising from the standpoint of spatial economics that per capita income differences among city-regions are so persistently high,

Michael Storper is in the Institut d’Etudes Politiques de Paris, 13 rue de l’Université, 75007 Paris, Paris, France; and Department of Geography and Environment, London School of Economics, Houghton Street, London, WC2A 2AE, United Kingdom. E-mail: michael.storper@sciences-po.fr and m.storper@lse.ac.uk.

especially in economies (such as the US) where there is a high level of interregional capital and labour mobility, and even within classes of structurally similar metropolitan areas. The barriers to income convergence are familiar to students of international development, where limits to capital and labour mobility can be very high. Once such limits are taken into account, development economics usually turns to differences in specialisation as the source of developmental differences, with such differences involving everything from comparative advantage to human capital differences, to institutions (Sala-i-Martin, 2002; Barro, 1996; Trefler, 1993).²

Along these lines, this paper argues that research on differences in metropolitan growth can draw lessons from international and comparative growth economics. The latter has recently built up a pragmatic view of growth differences as the product of three principal forces: specialisation, human capital and institutions. As in international economics, however, the principal challenges to explaining growth reside in the potentially complex chicken-and-egg relationships of these forces, as well as in their relationship to path-dependent evolutionary forces that can be set into motion by unique events or shocks. In the review that follows, we explore these forces in the metropolitan context and knit them together in a systematic way that brings us closer to a more adequate framework for explaining why cities grow differently.

2. Existing Approaches Have a Limited Ability to Explain Causes

At the root of differential incomes are differences in the composition of an economy, itself reflecting differences in factor content (proportions) and factor prices from one economy to another. Concretely, when a sector (or sub-sector, or group of tasks) is concentrated in a certain region, it leaves a strong imprint in terms of the quality of jobs

and, in the medium term, its growth rates will be highly correlated to overall local growth. As a recent indicator of this, Galbraith and Hale (2004) note that the rise in the share of national personal income of just *four* (out of about 3000) US counties in the late 1990s is sufficient to account for the majority of the increase in geographical intercounty income inequality in the US in the 1990s. These counties housed the core clusters of the US high-tech boom.

Most standard economic theory is not fully comfortable with the notion that 'specialisation matters' over the long run at the international level; by assuming a world without economies of scale (including agglomeration economies), it concludes that convergent factor proportions will ultimately 'wash out' whatever effects sectoral specialisation might otherwise have on national incomes (Krugman and Obstfeld, 1991). Economic geography, by contrast, has potent theories of why and when sectors will geographically concentrate and hence promote specialisation: these are the theories of agglomeration based on internal trade costs (sometimes called production 'sharing'), labour pooling and localised technological externalities (Duranton and Puga, 2004; Rosenthal and Strange, 2001). In a world with these features, the story is very different. Sectors or activities that are geographically concentrated (generating local specialisation) are at different points in their developmental cycles from those that are spread out. They are characterised by different factor proportions (labour- versus capital-intensive, for example) and require different qualities of factors (Trefler, 1993; Norton and Rees, 1979). Moreover, they have different terms of trade from the spread-out sectors. Generally speaking, geographical concentration and local specialisation are highly present in newer or more innovative activities, or those that depend on hard-to-imitate skills or knowledge. As a consequence, they can earn temporary rents on their outputs, which they

can then prolong through further innovation. Moreover, in these sectors not only is there a rent effect, but it is difficult for any supplier to have a decisive impact on prices, so the places that specialise do not undermine their own positions easily, as is the case for those places specialised in products with easily expandable supply (mature or standardised products).³ There are other sources of such rents, such as locationally specific non-tradeable goods and services (landscape, architecture) and possibly certain activities with very high internal scale economies, but these are likely to be weaker (subject to more substitution) than those based on knowledge, skill and innovation.

Cities can also be specialised in a way that makes them poorer than the average. As noted, agglomeration economies are tied, in one way or another, to the intermediate output structure of sectors: interfirm transactions; labour pooling; and technological spillovers. All of these could make their host areas richer than average, assuming that their productivity or rent-generating effects are greater than congestion costs. Yet one can imagine also that economic activities that have none of these locational processes find themselves together in a certain city or region simply because it has the right factor supply for that sector (say land, or labour or transport access). This form of 'concentration' reflects a simple comparative advantage, but not an agglomeration. In practice, then, not all high location quotients can be taken to indicate specialisation in the sense theorised by the New Economic Geography. The fast-growing cities in the US interior West have economies specialised in routine functions characteristic of the later phases of the product cycle, with generally high firm-level scale economies, but with few of the sharing, matching and learning features of agglomeration processes that generate local economic rents (Scott, 2008). The slower-growing cities of the Northeast have higher incomes and more specificity based on specialisation, with the attendant

benefits (Drennan, 2002).⁴ Ultimately, then, any framework for explaining urban growth and income differentials has to engage with the causes of specialisation.

Much research on why some city-regions have higher income levels or more overall growth than others does not tackle the question of the sources of income-differentiating specialisation, instead employing a method best described as inductive empiricism. Such research proposes a list of factors that are supposed to contribute to growth. A city is believed to grow by increasing its share of a sector's employment, or when its existing specialisations grow faster than national economy as a whole. A mix of factors such as labour costs, land costs, regulation, business climates and so on, are adduced as explanations for why the city does well or poorly in shaping the evolution of its economic base (see Glaeser and Shapiro, 2003; Glaeser *et al.*, 1992; Glaeser *et al.*, 1995). Impacts of specialisation are then assessed through export-base models and multipliers.

The principal omission of these approaches is that they make little attempt to identify causes. For example, there may well be a correlation between land costs and where an activity locates *within* a metropolitan region (expensive in the centre, cheaper in the 'periphery'). There is also a correlation between labour and land costs and which types of activities go to which types of cities, which mirrors international specialisation patterns, at a very general level, between high-wage/capital-intensive places and low-wage/labour-intensive places. Hence, these costs matter in explaining why New York, San Francisco and London have little durable goods manufacturing, while certain cities in the 'Deep South' of the USA or southern Europe have a lot of it. However, there is not a relationship to why an activity locates in one *particular* city or another *within* a class of structurally similar cities; for example, they cannot account for why San Francisco has so much more high-technology industry than Los Angeles.

Comparing productivity levels among cities does not get us much closer to a real explanation. High-cost city-regions have high relative productivity levels, but this is because they have different activities from low-cost city-regions, so there is the need to break into the circle of causes and effects (Sveikauskas, 1975). Thus, only for sectors that are generally deconcentrating (and hence are potentially footloose) does interurban productivity difference affect where they are subsequently located, but this relocation is generally to a different structural class of cities. Even so, it cannot tell us *which* cities in that other class of cities (lower-cost ones) will get the activity, nor can it disentangle whether productivity levels are causes or effects of such locational changes.

Recent work in urban economics is able to make finer distinctions than these approaches, by including a greater number of forces, doing so systematically and using an equilibrium framework to assess trade-offs among them. Thus, Glaeser (2007) proposes an integrative ‘three-equilibrium’ approach to comparative urban growth, in which incomes, population and specialisation are simultaneously determined. Population is determined through the ‘builder’s equilibrium’, through elasticities of housing and population, where housing stock changes are determined through the regulation of land development. Wages and employment result from the ‘firm’s equilibrium’, where productivity is driven by skills, but wages are driven by elasticities of employment in turn determined by population growth, from the builder’s equilibrium. Household preferences for a region’s amenities are elastic to nominal income (hence wages) and could affect wages as well through population pressure shaped by the stock of amenities. The interaction of these three processes thus yields structural determinants of growth in general and the positions of different types of cities according to how the three processes work out in specific cases. Broadly, in the contemporary US, there are ‘high-wage, high-skill, housing-scarce’ cities

and lower-wage, lower-skill, housing-abundant cities. This corresponds to the split mentioned earlier between the US urban Northeast and West Coast, on one hand, and the Sunbelt cities on the other (price-based, quantity-constrained growth versus quantity-based, price-constrained growth).⁵ Sub-classes are possible, when more detail about certain kinds of amenities (especially climate, consumption and leisure) are added.

How well does this approach do in sorting cities into different observable growth and income categories? Consider the comparison, again, of San Francisco and Los Angeles. Both are in the general class of high-wage, high-income metropolitan areas, where housing prices have increased at a higher rate than population in the past few decades. Whereas their per capita incomes converged between 1945 and 1970, they have diverged since, leading to a one-third difference today.⁶ The model suggests that such a difference could be generated by a different position in the builder’s equilibrium—i.e. if SF had less housing expansion than LA, thereby limiting population growth and driving wages up. Most empirical indices (with all their methodological problems)⁷ suggest that the overall quality-of-life amenities are ‘higher’ in SF than LA so, according to the model, workers should actually be prepared to take lower nominal incomes in SF, *ceteris paribus*; the builder’s equilibrium would need to be a very powerful force for income divergence. However, since 1970, the two regions have had roughly proportional population growth and roughly proportional growth in the number and size of housing units.⁸

The only candidate left for explanation, then, would be in the area of the economic base—i.e. the locational behaviour of firms. During this period, SF grew to be the centre of the world’s ICT economy and LA lost much of its previous high-tech base in aerospace. How can this be explained? According to the model, it should be the result of being at different

points in the builder's equilibrium, with San Francisco restricting growth and hence getting a more skilled migration stream. SF did in fact get a more skilled migration stream, but the builder's equilibrium cannot explain it: both regions doubled their population and housing stock. Thus, for such an approach to work, its three constituent equilibria would have to be based on realistic facts.

Explanation would then turn to something additional that determines worker productivity and possibly in-migration. The three-equilibrium model holds that this could be 'worker interactions' (Glaeser, 2007). Even if we grant, plausibly, that high-tech workers in SF augment their productivity through various kinds of interactions within the large regional labour pool, it cannot be used to explain the origins of the divergence, since such major interaction could have only occurred once a large pool of people in the electronics industry were already in place. Such a pool, once established, may in turn have many recursive effects on productivity and specialisation, but it has to start with some other force that is omitted from the model.

Such models, no matter how descriptively sophisticated they may be, fundamentally assume that 'jobs follow people' when it comes to household locational behaviour and that people follow jobs, when it comes to firms' locational behaviour, and that the two are joined at the hip through the endogenous and simultaneous determination of the housing stock as the source of population versus wage growth (Muth, 1971). The assumption of simultaneous movements towards all three equilibria possibly means that there are no directions of causality, nor time-dependent developmental sequences that might need to be considered. There is, in essence, no history to such models.

This lack of history or sequence is also canonised in the use of the 'spatial indifference assumption' in the firm's equilibrium. According to this assumption, firms go anywhere where they can maximise productivity,

in an infinite combination of trade-offs of such things as worker skills and wages, made possible through the perfect divisibility of production and malleability of factor proportions according to location. In this world, there can be no economies of scale, neither at firm level or at production system level (localisation economies). This approach is fundamentally incompatible with 'second nature geography'—i.e. a world where spatial concentration and concomitant specialisation are endemic (Fujita and Thisse, 2002). If scale economies exist, then decisions about where to locate production do not involve the perfect divisibility and trade-offs described by the firm's equilibrium; hence, the other trade-offs to housing, wages and amenities, would no longer be fully simultaneous because factor elasticities would be circumscribed by internal and external economies of scale. Thus, the interaction of the three constituent markets (builder's, firm's, household's) cannot be simultaneous; grasping their effects on development therefore requires models that have hierarchies of causes, directions of causality and temporal sequences.

What if we change course and assume that the spatial concentration of firms is the most important source of differential urban growth—i.e. that it is the starting-point for complex intertemporal sequences of causes and effects? The most prominent recent approach to spatial concentration of firms is the New Economic Geography model of Dixit, Stiglitz and Krugman (Krugman, 1991a). It tries to reconcile neo-classical assumptions with spatial concentration, by assuming that jobs go to people, who in turn search for the benefits of scale economies in consumption (cheaper goods, more variety) by moving to large urban 'home markets'. At the same time, these people also go to jobs, because firms concentrate at the centre of these large markets, thus allowing them to produce more efficiently with economies of scale. So the causes of urban growth are said

to be fully simultaneously two-directional (Krugman, 1991a). Does this give us the tool we need to understand income and growth differentials?

Undoubtedly, it captures one source of spatial concentration today—the city as a centre of diversified, major consumption. In this model, cities below the optimum size thresholds would suffer a growth penalty and those close to them would grow the most. It could also capture certain specificities: for each class of goods, size thresholds exist and those cities closest to these sizes would grow according to the growth of demand for the goods corresponding to their market size, in a sort of updating of central place theory. Another source of differentiation among cities would therefore come from the possible clustering of supply chains for different goods, if trade costs are strongly positive.

Yet all in all, it probably does not give us very strong tools for understanding different rates and types of growth among cities in similar structural classes. One reason is empirical: for most standardised goods and services, the trade costs of linking production to retail in an advanced economy are so low that there is relatively little local backward linkage from the local home market to the input supply structure. Another reason is that any such clustering together of the upstream, intermediate, activities in a production chain, even when present, is rarely caused by serving a local final market. Thus, while New York is specialised in financial services because its intermediate producers concentrate together to serve a large concentration of banks and brokerage houses, these latter are not concentrated in New York City because of its large population and the ‘jobs–people–jobs’ story of the Krugman model.

In the end, the question of the origins and growth of specialised localisations of economic activity remains largely outside all these ‘new urbanisation economies’ approaches.

3. The Sources of Specialisation

The origins of economic specialisation are in some ways the elusive ‘holy grail’ of development economics. In comparative advantage approaches, there are almost always many places with similar comparative advantages, so that we have difficulty explaining why a particular set of places gets the activity in question without resorting to additional analytics. Agglomeration economics aspires to account for why, once an industry gets launched in a place, it tends to keep growing for a long time, with possible lock-in effects. Economists have called this the ‘history matters’ part of the growth process (Krugman, 1991b) and it has been incorporated as well in evolutionary economic geography (Boschma and Martin, 2007). Once this is done, we are no longer in the realm of comparative advantage theory. Yet nor does agglomeration economics provide us with general explanations for why such agglomerations get started in the first place and the literature usually falls back on ‘unique’ historical circumstances to explain each agglomeration in an *ad hoc*, case-by-case way. (Stanford University for high-tech in the San Francisco area, or good weather for entertainment in LA, for example), on the one hand, or excessively general explanations such as business climate or factor costs, on the other.

It is indeed conceivable that specialisation is largely *accidental*: being in the right place at the right time attracts a sector to a place. In this view, then, the causes of specialisation are external or ‘exogenous’. One can think of a combination of these two latter views: thus, accidents or unique features of a region—such as its historical experience, unique skills or unique founding events (a key entrepreneur or technology) generate initial ‘seeds’ of specialisation. Such seeds, however, are often planted in at least several places. Powell *et al.* (forthcoming) show that about 10 US regions were early patenters in bio-tech/life-sciences

industries, but only three have emerged as agglomerations. So we need to know about both the planting of the seeds and how they are subsequently cultivated.

Moreover, as is suggested in a highly simplified way by product cycle models, there is a moment in the life of most industries when their internal trade costs decline to the point that no local measures can suffice to maintain their core agglomerations (Norton and Rees, 1979). Regional economies appear to react quite differently to such evolution. In some cases, it leads to major job loss. In others, the inevitable loss of routine production activity is compensated by moving ‘up’ the product ladder or generating new activities through innovative entrepreneurship within the region (Saxenian, 1994; Amsden, 1989). All urban economies are therefore faced with the dilemma of how to adjust to these external shocks. Successful adjustment comes essentially through sectoral succession or innovation: capturing activities that can become the basis of new regional economic specialisations, or retaining the retainable parts of existing specialisations by reinforcing comparative advantage in a certain part of the activity’s value chain, or by moving up the product quality ladder within an industry already present in the region. Explaining this

adjustment process, once again, depends on our ability to explain the origins of renewed specialisation, and innovation that moves a regional economy up the product/price hierarchy.

If we apply these lines of thinking to metropolitan economies, the task then becomes two-fold. First, is there some way to explain the origins of successful specialisation among metropolitan regions in a similar structural class of regions? Secondly, given the inevitability of change in the locational patterns of sectors in such a region’s economic base, why do some places seem to do better at adjustment than others, effectively getting from one pattern of specialisation to another over time? Comparative growth theory, mostly as applied to international growth and development comparisons over the medium- to long-run, offers some hints. For the question of origins, it suggests that we look at either human capital or institutions. For the question of succession and change, it suggests that we look again at how institutions affect the *adaptive capacity* of economies (Rodrik, 2007; Hausmann and Rodrik, 2003; Pritchett, 1997).

These endogeneity issues with respect to regional specialisation can be visualised as indicated in Figure 1.

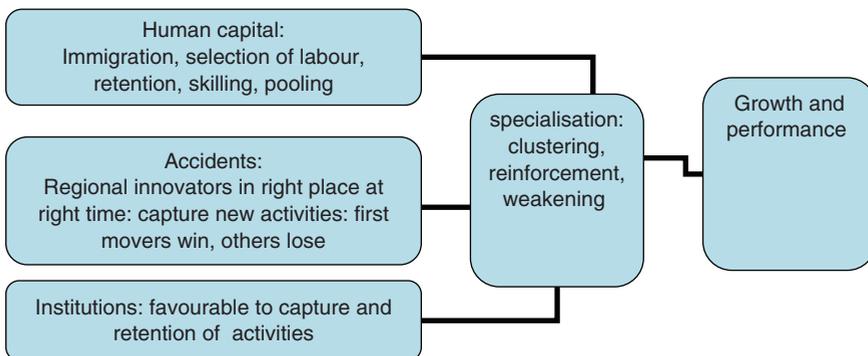


Figure 1. What might cause specialisation?

Let us now, therefore, turn to the feedbacks between specialisation, human capital, and institutions.

4. Human Capital and its Origins

It has been famously observed by Lucas (1988) that skilled people congregate in expensive cities to be near other skilled people. Many human capital models of regional development extend this idea, implicitly, to the premise that 'jobs follow people'. Specialisation of a metropolitan economy would be driven by the characteristics of its labour force, building labour quality into the standard notion that specialisation responds to factor endowments. A further step is to argue that, since capital is usually more portable than labour, human capital becomes the factor determining specialisation, in some relevant time-frame.

For this observation to become a complete 'jobs to people' argument, however, the determinants of the location of the labour force have to be identified. To account for intercity income differences, we would want to know why some expensive cities do better at attracting skilled labour than others and what triggers the divide between more- and less-skilled cities. This is even more a mystery when we remember that regions within a country are highly open to interregional migration and somewhat open to international flows of people. This means that the stock of knowledge of any given city-region is intimately related to national education, R&D and labour migration between regions.

For such differences to persist, then, the regional human capital stock at any given point in time must be, in part, *caused by* the regional characteristics that attract, retain and repel people with different kinds of skills. For some analysts, these characteristics are amenities, which are held to drive residential choice,

which in turn (through the jobs to people mechanism) drives specialisation (Florida, 2002; Glaeser *et al.*, 2001). In the creative cities framework, specialisation is driven by the attraction of creative workers (whose main component is highly educated workers, with a high proportion working in high technology and finance). These workers in turn are said to accumulate in places because of the amenity of 'tolerance', which is operationalised through the composite variable 'diversity'. The 'amenity city' argument generalises this to both highly educated populations (high-culture amenities and bohemian amenities) and less-educated populations (sun, low density).

The problem with all human-capital-driven regional economic growth models is, once again, directions of causality. Skilled people appear in most cases to precede the creation of amenities, not principally to follow them. Instead, household preferences for residential amenities may indeed drive specific locational choices *within* metropolitan regions, but not *between* them, so they generate intrametropolitan sorting but not regional development as a whole (Cheshire and Sheppard, 1995; Oates, 1969; Hilber and Mayer, 2004). Moreover, since skill groups are sorted among cities according to different specialisations (finance workers in New York, directors and producers in Los Angeles), it becomes difficult to prioritise the role of employment opportunities as opposed to amenities, since amenities are a highly aggregated and much less geographically distinct category than employment (for example, many places have restaurants and art galleries, few have high-level investment banking industries).

As noted, some models emphasise that skilled workers seek interactions with other skilled workers, which in turn induces a basic split between skilled (high-wage) cities and less-skilled (lower-wage) urban economies. Yet it is hard to nail down the

causal force of workers seeking interactions. It may indeed play a role in reference to an already-specialised city-region, which then attracts individuals with specialised skills in those activities. It is more difficult to see how it could be an original cause, since without there having been previous rounds of development that led to the concentration of skilled workers in a place, why would others with specialised skills want to migrate in?

In the now-classical ‘matching’ model of labour force agglomeration, workers are said to concentrate together when they work in a high-turnover industry, because they want to be near a large pool of potential employers, as this raises the probability of securing a new post following layoff; and employers cluster together in order to tap into a large pool of available labour, avoiding hoarding during downturns (Jayet, 1983; Combes and Duranton, 2006). Such a model is descriptively persuasive, but once again it seems as if it should be an outcome, rather than an original cause of a certain form of local economic development. Thus, the clustering together of many such firms with unstable labour demands might generate the regional labour pool through in-migration or regional learning-by-doing, and then the two become mutually reinforcing causes (Scott and Storper, 1987). Moreover, nothing in these models enables us to understand why some cities/regions become specialised in those parts of the economy that have high labour turnover versus those that are in routine production with longer-term employment relations. Nor does it provide an explanation of why cities get pools of labour that match in the particular sectors they do. Matching in the high-turnover garment industry generates a labour pool that is not a pool for the high-turnover Internet services industry. And of course, each of these pools of labour receives very different wages, including variable wage premia for both high turnover and interaction.

Thus, while it is evident that human capital must be an important influence on the course of growth of all urban economies, and that differences in labour force appear to be associated with differences in urban growth and income levels, surprisingly little has been established about the causes of differences in human capital, nor about whether it is human capital that drives patterns or specialisation, or the other way around, or even whether still other forces drive both specialisation and human capital—and, hence, income—differences, within complex intertemporal sequences of developmental processes (Rosen, 1983).

A final candidate for the deep causes of differences in human capital might be our other two forces, institutions and specialisation. All of the mechanisms of human capital creation—foreign and domestic migration and training and retention *in situ*—could be fundamentally affected by initial specialisation. Historically, relationships tend to deepen between the local economic base and various kinds of training institutions. Specialisation, by initially attracting both domestic and foreign migrants, could then set up migration ‘chains’ in communities and regions of origin, that become self-reinforcing through signalling. Then, as workers learn about how other workers improve their skills through interaction, they might migrate in. Stated thus, we are also making a potential argument about institutions in the receiving region and notably about mutual causality between specialisation and institutions that sets off development of the labour market. Along these lines, if tolerance and openness to newcomers are ‘amenities’ that attract the skilled, then they are fundamentally institutional in nature; rather than a pure ‘jobs to people’ argument, the causality runs from institutions to people, or possibly from specialisation to institutions to people.

A picture of these endogeneity issues with respect to labour supply may be seen in Figure 2.

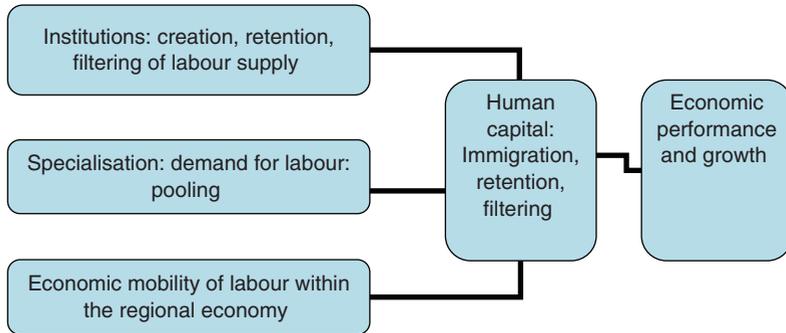


Figure 2. The sources of regional human capital.

5. Do Institutions Select for Human Capital and Specialisation over Time?

A third major branch of international growth theory argues that *institutions* are the ultimate shapers of long-run economic growth (Rodrik *et al.*, 2004; see also Glaeser *et al.*, 2004; Acemoglu *et al.*, 2004). By ‘institutions’ is meant a variety of things, ranging from the ways the formal *de jure* rules of political institutions affect their efficiency in facilitating economic activity, to what we might call *de facto* governance, referring to the real, on-the-ground ways that public-sector agencies and private-sector groups and individuals interact in detailed ways to shape the rules and resources of the economy, and the ways they mobilise beliefs and norms in so doing (North, 2005).

There is strong reason to believe that appropriately constructed institutionalist arguments and evidence could advance our ability to explain both why certain places capture certain activities and also why they manage development transitions better than other city-regions. Three fundamental areas of institutional performance (*de facto* and *de jure* combined) are at the centre of growth theory: the ways that institutions shape the microeconomic environment (including what is commonly called the ‘business climate’,

covering such things as confidence and the ways it affects transactions, discounting and investment levels); the ways that they shape labour force participation and effort levels (sometimes known as the ‘social policy environment’); and the ways that they shape problem-solving, which determines how well the economy captures new opportunities or misses them (how it adjusts to changing technologies and competitors). There is no precise institutional blueprint for these features; rather, they represent outcomes for which there are many functional equivalents, depending on the context (Rodrik, 2007).

In the overall puzzle of explaining regional economic development, the area of institutions is perhaps the most complex and least well-explored space. It has been somewhat easier to make progress on it when investigating comparative national growth processes, because national boundaries describe relatively distinct *de jure* institutional structures and national data on politics and legal structures, as well as international survey data, allow for comparison (Djankov *et al.*, 2003; Hall and Soskice, 2001). Such data are very limited when it comes to the metropolitan scale; moreover, *de jure* institutions are highly subsidiary to other levels in most countries and political action networks are highly open to other regions, blurring their ‘regionality’. Therefore, the empirical record is much

sparser. Still, by extending arguments from the institutionalist literature to the city-region level, we can ask how institutions might affect regional specialisation, incomes and growth.

5.1 The Formal Structure of Institutions and the Regional Policy Process

The most obvious formal power that localities and (in many places) regional authorities have to affect economic development is in the regulation of land use and in major public investments, such as infrastructure. Land use and public investment decisions are thought by some economists to have strong effects on the conditions for labour market/business network formation and sustenance, and hence on specialisation. For example, some economists believe that greater restrictions on land use will drive up housing prices and in turn have a selection effect on the labour force (selecting for high-skill, high-wage labour) which would, in turn, price in certain specialisations and exclude others (Saks, 2007). Yet many questions remain. One is whether the empirical indices of regulation used by this literature are of sufficient quality to sustain arguments about its importance.⁹ A second is whether high-wage regions really consistently grow more slowly than low-wage regions and, if so, whether this is *because* of regulation of the housing stock. Another question is whether regulation significantly affects overall regional housing stock changes, or whether it principally affects intraregional distribution of changes in housing stock (since in reality regulation is almost all municipal). Still another is whether preferences for housing type are so rigid that there is a high elasticity of housing stock to employment; stated another way, that such preferences override the choice of, say, smaller and denser housing in high-price areas (or bigger housing towards the metro fringe in these areas) and push people, *and with them*, whole groups of jobs (specialisation) to other metropolitan regions. In other words, for land use controls to be

shown to be the ‘dog that wags the tail’ of the regional economy (specialisation, wages and demographics) requires very strong assumptions and strong evidence. We remarked earlier that the case of San Francisco, the American high-income city-region *par excellence*, does not lend much credence to the overall argument; the Bay Area has doubled its housing stock and population and, at the same time, it has ascended to the top of the world metropolitan income hierarchy, since 1970. Also as noted earlier, an alternative literature holds that the dog that wags the tail of housing prices is labour productivity and wages, which, in a world of factor proportions that are not fully elastic due to economies of scale, is a consequence of urban economic specialisation (van Nieuwerburgh and Weill, 2009).

A different approach to formal institutions concerns the way they might shape the regional political process more generally and especially ways political attention is shaped and policy agendas established (Jones and Baumgartner, 2005). Political economists have recently theorised that there are trade-offs between the efficiencies that can be gained from greater size of jurisdictions, and the losses generated (Alesina and Spolaore, 2003; Aghion *et al.*, 2005). The bigger the unit (for example, a city, a region, a nation), the more likelihood there is that there are more heterogeneous preferences of the people within it and hence the likelihood that many of those preferences will get ‘washed out’ in the conflicts and compromises that must take place in big jurisdictions, or at least will involve higher transaction costs for decision-making. In international growth studies, the performance of countries can be partially attributed to how successfully they combine the advantages of scale while enjoying sufficient convergence of preferences¹⁰ to be able to make strong decisions that have public support (Alesina and Spolaore, 2005). Such differences in the distribution of formal governmental competencies within metropolitan regions do receive

attention in some surveys of local and regional governments, but there has been little effort on the part of scholars to determine whether the resulting size structure of jurisdictions might influence voice, preference aggregation, coalition formation and decision-making (or definition of policy agendas) in different metropolitan areas (Carruthers, 2002; Kenyon and Kincaid, 1991; Kenyon, 1997; Peterson, 1981). The notion that this *de jure* structure might give metropolitan areas different capacities to shape their economies has received implicit attention in a certain applied literature that argues that metropolitan-scale decision-making is required for effective co-ordination of a metropolitan-scale economy—a literature which generally concludes, in contrast to the Tiebout (public choice) literature, that more centralisation might be a good thing for helping city-regions to capture and maintain growth (Rose-Ackerman, 1983).

These theoretical insights suggest a number of ways that the formal structure of jurisdictions might influence policies that affect metropolitan economic development. In a region with many small jurisdictions, localities have more homogeneous preferences and lower costs of debate and compromise, allowing more initiatives to see the light of day; however, those initiatives that require areas will only be able to be implemented by cross-jurisdictional coalition-building. Regions with less fragmentation (bigger, more internally heterogeneous jurisdictions) may have problems in reaching decisions and hence in generating novelty, but once decisions are made they can more easily implement large projects. This reasoning leads to the hypothesis that the ‘worst of all worlds’ in terms of the *de jure* structure of decision-making would be to have neither the size that promotes consensus, nor sufficient centralisation to implement large regional projects (see Rose-Ackerman, 1983; Stiglitz, 1983). This type of formal structure can be expected to generate stand-off or blockage (see Figure 3).

We have considered a simplified picture of the geography of formal decision-making institutions. Many of them, especially the larger cities in a region, have complex *internal* jurisdictional structures. As shown by Aghion *et al.* (2005), as cities become both larger and internally more heterogeneous, the propensity to divide political authority into districts or wards increases. The ability of a big unit, then, to exercise the ‘power’ of its size in a region will depend in part on how it organises its internal decision-making. This in turn depends in part on its municipal form of government (strong versus weak mayor), the correspondence of districts to different preference structures and the organisation of voting coalitions within that city. The formal institutional structure of a city-region must, in any case, be considered a multiscale patchwork of jurisdictions, decision-making processes, scales and capacities. Seen in this light, its complexity is indeed intimidating to the researcher, yet it offers promising insights into the formation of regional policy agendas and the pattern of decision-making.

5.2 Actor Networks and the Informal Sources of Institutional Functioning

Like different countries, city-regions often manifest strong differences in their inherited (*de facto*) informal political cultures and forms of political mobilisation. These could plausibly contribute to differences in governance outcomes. Some city-regions have longer traditions of intense community action and established patterns of government–business–community co-operation there, while others have much more top-down political cultures, for example (Logan and Molotch, 1987). ‘Social capital’ indices, that measure such patterns of participation, show that there is much more participation in some city-regions than in others, although the meaning of this for political outcomes has not yet been established by research (Putnam, 2000).

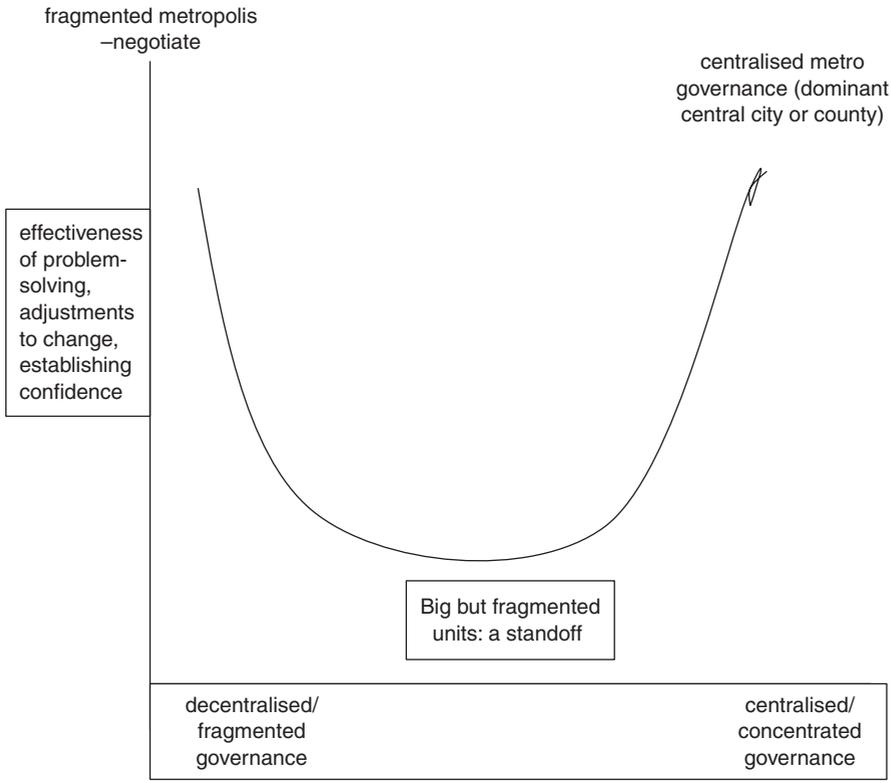


Figure 3. The formal structure of metropolitan institutions.

Regional politics that affect economic development might, by contrast, be structured less by the broad outlines of political mobilisation than narrower sectoral interests. In this case, specialisation has a strong influence on politics, in that specific business élites may shape political markets for ideas, as well as the policy formation process in a region (Dahl, 1961/2005; Cox, 1993). Powerful landed interests may also shape the choices made about urban development, especially in land use, through the place-based politics of landowner and developer groups (urban growth regimes) (Molotch, 1976). In this way, sectional preferences can find broad expression through lobbying, interest-peddling and other means of dominating the resource-allocation and policy-setting processes. When the concerns of these two literatures are

brought together, they suggest the interesting question of how the broad institutional structure and political process of a region interact with its sector-specific interest-groups and institutions, such as business associations and labour market networks (Jaher, 1984).

The conception of institutions that we propose to import from growth theory is thus broader than in most of the urban politics literature (Persson and Tabellini, 2006; Rodrik *et al.*, 2004; Glaeser *et al.*, 2004; Jones and Baumgartner, 2005).¹¹ In addition to focusing on élite processes directed to the deliberate use of public power to extract rents and influence land development, it also emphasises complex, dispersed collective action problems such as how actor networks are formed, supported and eliminated, and how this affects policy (Padgett and McLean, 2006; Powell *et al.*,

2005). Such actor networks might be those of specific groups, or they might be cross-cutting. In the regional development literature, analysts have been interested in the *sector-specific institutions* that make a region able to help a particular industry to flourish. This is especially the case with respect to high-technology clusters and flexible production networks and ‘innovation-based’ sectors (Saxenian, 1994; Becattini, 1990). Questions about how production networks are co-ordinated, moral hazards contained and transactions costs minimised, as well as whether sector-specific public goods are provided, are the object of a vibrant literature (Scott, 1993).

Economic sociologists have investigated actor networks at several scales—sectors, professions and labour markets (Powell *et al.*, 2005; Granovetter, 1995). Informal urban institutions could, in this vein, be thought of as actor networks at two criss-crossing scales. These scales draw from the classical sociological distinction of ‘community’ and ‘society’, and their doubles, ‘bonding’ and ‘bridging’. The actor networks of various kinds of communities—sectoral, professional, neighbourhood, ethnic and so on—constitute the basic tissue of political life in a city-region, but

how these networks interact through ‘bridges’, constitutes the broader ‘society’ of the urban region in question and the coalitions that matter for shaping economic development (or blocking it) (Storper, 2005).¹² In the wider growth framework proposed here, we need to ask how such networks are formed and changed, by forces such as changes in the regional human labour force and specialisation. Hence, the directions of causality are likely to be multiple and complex. These endogeneity issues with respect to institutions are summarised in Figure 4.

6. Events, Accidents or Structural Factors?

Research on economic growth has long struggled with the relationship between structures and events. Formal economic models, as well as inductively empirical approaches, are structural, in the sense that they seek the parameters that are associated with growth across a wide panel of cities. Yet decisive events are often cited when the medium-run fates of particular economies are cited: Detroit declines ‘because’ of the restructuring and decline of the American car industry;

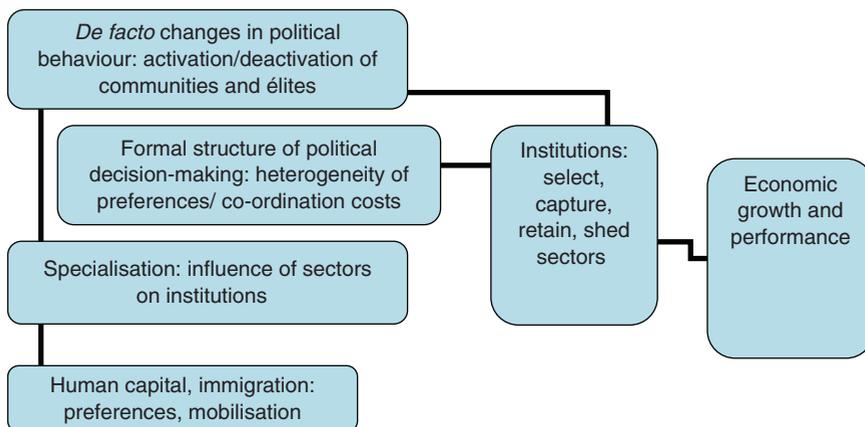


Figure 4. The sources of institutions.

Los Angeles has declining per capita income 'because' of a massive wave of low-skilled immigrants in the 1980s; Washington, DC, becomes richer 'because' of rapid growth in the high-wage lobbying industry. Often, the latter are treated as exogenous shocks in formal models, but if such events are frequent and important shapers of growth, then much that is important is simply being avoided.

This ambiguity is particularly potent with respect to agglomeration-generated specialisation. Theories of agglomeration, in effect, show that localisation economies have a structural character in that, once started, they can be strongly self-reinforcing; but, as noted, explaining their origins may require recourse to one-off events as well as structural factors. In the framework we have developed in this paper, therefore, human capital and institutions can be situated clearly as structures, but specialisation may be both structural and event-driven.

The spectacular example of Silicon Valley provides a perfect illustration of these issues. In the late 1950s, it was not clear that an area south of San Francisco would ultimately become the world centre of information technology and propel the entire metropolitan area up the US income ladder. The initial pattern of semi-conductor production in the US was quite dispersed and, if there was any nascent geographical centre of gravity for the sector, it was the Northeast coast, from Boston to New Jersey; the principal forerunners of the IT industry—i.e. the radio, telephone and television equipment industry—were all located in that region. However, because the semi-conductor industry was extremely new, it lacked a clear 'production process' and 'commodity chain' and therefore did not have clearly defined 'factor demands' or even linkage patterns. There were a number of areas in the US with abundant engineers working on what would become the new technology, from northern California to southern California, to the east coast and even in the south-west.

This is what we previously labelled a 'window of locational opportunity' (Scott and Storper, 1987). Such moments of openness existed in the late 1920s with the American aircraft industry, or the film industry in the 1900s. In the life sciences industries, such a window appears just now to be closing around the centres of Boston, San Francisco and San Diego (Powell *et al.* forthcoming).

The problem is that there are often *many places* that can satisfy the initial factor requirements of an industry. Perhaps in the Silicon Valley case, there *were* many regions that were—more or less—equally well prepared to become the future world centre of the IT industries. Yet only one of these 'candidate regions' happened to get ahead just a little earlier than the others. Is this because some actor in that place happened to be the first to come up with a break-through 'killer application' that tipped the agglomeration economies towards that place: by taking market share, suppliers streamed into the Valley to fulfil new needs, network efficiencies grew and other places found themselves out-distanced, even though they were 'about equally good as one another' prior to this tipping-point? The extreme version of this story of accidents or idiosyncratic causes holds that William Shockley, the inventor of the chip, moved to Silicon Valley because he wanted to be closer to his mother, who lived in Menlo Park. This then locked in the locational structure of the industry. This kind of accidents-of-history explanation is now incorporated in the New Economic Geography's basic core-periphery model (Fuchs and Shapira, 2005). It fits well the case of the aircraft industry in the 1930s, where many localities in the US or elsewhere were equally well placed to become major centres of aircraft production, but Los Angeles got ahead because of a single event: Donald Douglass invented the DC-3 in Santa Monica and it subsequently captured a very high market share of commercial aviation. Its produc-

tion system in southern California expanded as a consequence and this tipped the geography of aircraft production there, by attracting in other producers who wanted access to its supply structures and its rich community of knowledgeable experimenters (Scott, 1993). The idea that unique events powerfully influence why some cities grow and develop one way versus another, in the medium run of 30–50 years is a difficult pill to swallow for those who formulate urban policy and want to predict the outcomes of their efforts and expenditures, because it implies that there is a chance element in economic development (Boschma and Kloosterman, 2005; Rigby and Essletzbichler, 1997).¹³

Do these events deserve the analytical status of ‘fully exogenous accidents’? It may be that certain pre-conditions must be satisfied even to have the possibility of the right breakthrough occurring and then locking in the geography of the sector. If this is true, then there is a regularity that can be explained by social science. For high technology, there is a long laundry list of such pre-conditions argued to have been necessary—presence of universities, certain pools of engineers and so on. A different laundry list is used for financial centres, another for ‘creative’ cities and so on, just as in the past such lists were developed for manufacturing centres in the developed countries. Yet it is a regularity more akin to ‘climate’ than to ‘today’s weather’: if pre-conditions exist in many places, but only some develop, then there is a wedge between broad and deep structural conditions and outcomes (Hodgson, 1993). It subsequently becomes difficult to estimate the structural parameters linking such pre-conditions to precise outcomes.

Such regularities might be more related to institutional context than to a precise list of factors. There are many examples in the history of innovation where superior ideas do not get implemented because they do not find such a favourable environment (Mokyr,

1990; North, 2005). Thus, if potential seeds of agglomeration emerge in different city-regions, they may flourish in some, but wither on the vine in others. Some noted analysts of the Silicon Valley case (for example, Saxenian, 1994) argue that it was Silicon Valley’s institutions—the actor networks of the merchant semi-conductor producers and their culture of ‘open competition’—that caused it to capture the industry and that this generated the specialisation it still enjoys today as the county with the highest per capita income in the world’s wealthiest economy. Yet once again, we fall into the endogeneity hole: Boston and Phoenix were early centres of semi-conductor production and only later on did Silicon Valley pull ahead of them. These actor networks cannot be said to have ‘caused’ the industry to emerge in Silicon Valley, since they emerged contemporaneously with it, as a part of the growth of the industry itself. This would also be true of the labour force. The labour force could not have ‘caused’ the IT industry to concentrate in Silicon Valley because, in the 1960s, computer engineering was just a loose set of individuals, not a consolidated academic discipline. It is an endogenous outcome of the industry’s development, both as a category of skills and training, and as a geographical concentration of those people in Silicon Valley.

Two of the key elements of the Silicon Valley *de facto* institutional context today—the venture capital industry and the legal firms that specialise in technology law—are obvious products of the Valley’s development; they were not there prior to the existence of the IT cluster. The alternative institutionalist explanation holds that key actors in Stanford University (especially the dean of the Stanford Business School) established the ‘Stanford business park’ and, through this ‘institutional entrepreneurship’, allowed Silicon Valley to do better than Boston or Phoenix (Saxenian, 2000; Kenney, 2000). In this version of things, Motorola’s massive early investments in

Phoenix did not find a favourable institutional environment there, or Motorola simply made the wrong strategic choices, and hence the Phoenix agglomeration folded. Yet this would not seem to be the case for Boston, which had multiple and diverse actors who saw what was happening and tried to capture the IT industries at the same time Silicon Valley was also doing so. Saxenian's (1994) argument that they did not do so very well in Boston because the institutionalised corporate culture there was not well adapted to the merchant semiconductor sector is convincing, but it does not answer the question of whether the institutions in Silicon Valley were chicken or egg.

In the case of first-mover advantages of the type we are considering here, these conditions are likely to come from institutions, but not of the sector-specific type studied by Saxenian (1994) and many others. These places should have institutions that overcome existing problems—including existing interest-group practices for extracting rents, dominating perceptions or blocking other groups from getting attention in political markets and labour markets.¹⁴ Turner (2006) goes so far as to suggest that Silicon Valley benefited from the Bay Area's anti-establishment counter-culture in distinctive and important ways, mixing it with engineering and business élites to produce a new amalgam in a new edge city.

Long-term processes of economic development are not, thankfully, entirely dominated by first-mover advantages. The economy also affords abundant 'second-mover' opportunities. As sectors mature, they develop more complex internal divisions of labour and the trade costs internal to the sector tend to decline. This leads to the possibility of geographical fragmentation of the sector, and deagglomeration of some of its activities. Perhaps more importantly, as an industry matures, it develops a wide variety of product outputs, as well as many complex intermediate stages of production. Product differentiation and quality ladders are a basis for

interplace differentiation and competition, serving as formidable opportunity-creating devices (Grossman and Helpman, 1991). This leads typically to the establishment of secondary clusters.

Moreover, in order to prosper over the medium run, city-regions need to do more than capture first- and second-mover opportunities. They must adapt to two other challenges. On the one hand, they may attempt to retain existing activities in the face of competition by continuing to modernise them and, on the other, they must ultimately cope with the loss of certain activities. Such loss is inevitable when the evolution of organisation and trade costs in a sector eliminates the options of retaining the industry at realistic ranges of regional factor prices (greater fragmentation of the production process or maturation through the product cycle). In this case, sectoral succession through first- and second-mover specialisation must replace the losses and both are underpinned by local innovation capacities.

International development economics suggests that institutions create the conditions within which regional economic actors can engage in these processes of capturing second-mover activities, modernisation and sectoral succession (Rodrik, 2007). One needs only to look at the success of the Japanese automobile industry today to see how important second-mover strategies can be to the economic geography of development and how specific strategies do not transfer readily from place to place (Cusumano, 1985; Ellison *et al.*, 2002).

How might institutions at local or regional scale shape the 'action systems' that seem to underpin specialisation processes of capture, imitation, retention and adjustment/succession? By the time an activity is ready for secondary cluster expansion or challenger clusters, some systematic lessons about the sector in question will have emerged to serve as guides for second-mover regions. The institutions of certain places may allow them to form political coalitions that

make them able to satisfy these sector-specific conditions for establishing a secondary cluster. In this way, economic development becomes less arbitrary and accidental than for first-mover agglomerations. These will usually be elite actor networks in concert with other groups.

The labour market, as an actor network system, might have a structural role to play in capturing second-mover processes, sectoral succession and innovation-based adjustment. The formation of human capital in a specific sector in the economy is a complex phenomenon that has been analysed extensively in economic sociology. To simplify the argument: effective performance depends on skills, whether acquired on the job or in educational institutions. Yet acquiring and using skills also often depends on relationships—i.e. knowing where and with whom to acquire experience and develop further on the basis of existing skills (Granovetter, 1995). Relationships also link individuals to opportunities, through networks. Networks also allow for circulation of talent and for exchange of information that continuously improves skills. Some networks have a strong interpersonal basis, or a combination of impersonal and interpersonal connections between individuals. People in the New York financial services cluster generally have very different networks from those in the Silicon Valley information technology cluster. Thus, if we consider high-technology business networks, there are roots in institutions that pre-existed Silicon Valley, such as the military-industrial complex and the national university and R&D system, as well as private companies in the predecessors to high-tech, such as the radio and television equipment industries. It is difficult to imagine the perpetuation of specialisation in ICT in Silicon Valley, finance in New York, entertainment in Hollywood, or fashion in Paris or Milan without these sector-specific business and labour market networks, deeply intertwined with educational networks, that attract, convey skills to and retain the people

that are key to entrepreneurship and action in a particular industry.

However, path dependencies cannot explain everything: Silicon Valley is the exemplar of spatial and organisational *rupture* with the past. The case of the venture capital industry in Silicon Valley seems to express both the strength of a localised network—venture capitalists repeatedly claim that they do not want to go farther than a day's drive, so that they can regularly have face-to-face contact with the firms they run—and its initial establishment in a place that was far from the pre-existing technology heartland of the US in the 1950s. Another example of rupture is when, in the 1950s, New York took over a leading role in art from Paris, but the 'external shock' of the Second World War may have been the key to allowing New York's well-developed actor networks to seize this fortuitous opportunity.

7. Conclusion

The question of what makes a city-region grow remains a major enigma for research. The purpose of the preceding discussion has been to lay out the principal lines of causality that are candidates for major influences on why a city-region grows.

To begin with, the origins of such specialisation remain elusive to researchers. We showed that the principal candidates for explaining such origins are: some set of structural (sectoral composition/experience) pre-conditions for a certain type of activity; appropriate labour force characteristics; institutions; and, accidents or unique break-through events. However, the directions of causality between specialisation and labour force/human capital are not clear and it is plausible that they could reverse in complex recursive developmental feedbacks. It is also conceivable that specialisation is accidental at its origins, but subsequently structural, in the sense of being path dependent.

Moreover, both specialisation and the formation and evolution of the labour force should be influenced by institutions and policies, but we have little concrete evidence of precisely how this works at the scale of the city-region. Institutions might also be outcomes of specialisation or labour force characteristics, in the sense that sectors and labour markets are complex actor networks that could take the form of political or institutionalised forces in a region. Institutions might serve as the key forces that open up, or block, potential pathways of specialisation and labour force development. The principal point of this part of our analysis was to establish that no effective answer can be provided to the question of why some cities grow more than others without making much greater progress on the 'chicken-and-egg' basis of the origins of the regional economic base. A set of specific foci for analysis as suggested in this paper offers an agenda for going forward in this area.

The term 'institutions' covers a huge ground in political economics generally and this is even more so the case at the city-region scale. We suggested that it is useful to think of institutions as either 'general-purpose' in nature, those that secure the overall conditions for functioning of the economy, or 'specific' to a domain of economic action (such as a sector, cluster of activities or professional field). In the former case, such institutions are likely to be formal institutions of government, whereas in the latter they are likely to consist only sometimes of formal organisations, and many times to involve informal actor networks and coalitions.

In both areas, there are major gaps in our understanding of how institutions affect the growth of cities and in how to conceptualise and measure them. There are virtually no large-area rigorous data on how local authorities act to affect comparative advantage and specialisation and hence regional economic growth. Although there are many actions taken

by such authorities with the label 'economic development', there are few ways to establish their impacts or standardise what we mean by them. This is in part because only some of them take the form of laws or ordinances enacted; they also emerge in reports, in budget items and in other actions. The development of good measures and data would enable us to establish whether what they do has a major influence on growth or not. It is urgent to do this in order to get beyond the anecdotal cast of the literature on this subject.

Moreover, the formal institutions of city-regions rarely have clear sovereign structures, such as the central governments that exist for states. It is therefore difficult to speak of 'varieties of city capitalism' in the same way that this is considered in comparative political economy at the international level (Hall and Soskice, 2001). At the city-region level, we are instead faced with a patchwork of competing and co-operating authorities. In this paper, we suggested some lines of a framework analysing the performance of these nebulae, drawing from literatures on the size and internal preference structures of units and including co-ordination costs and public choice dynamics.

The analysis of informal institutions can take some inspiration from existing literatures on urban politics, as well as those on political regimes and comparative development. We argued, however, that the notion of informal institutions could usefully be expanded to include many different kinds of actor networks in the regional economy. In spite of strong theoretical reasons to believe that they are important, there is a woeful lack of rigorous empirical measures of them at city-region level, in a way that would enable us to establish their contribution to growth.

7.1 What to Do? Two Strategies

Two future strategies to analyse the causes of growth of cities can now be suggested. The first would consist of structured, in-depth

comparisons of the growth trajectories of different cities, using a set of standardised categories representing the issues identified in this article. The result would be rich in detail and 'on the ground' accuracy, but it would have few degrees of freedom.¹⁵

A second approach would require large-area data. As noted, the problem with most existing econometric analysis of urban growth is that it may identify certain structural determinants of growth, but does so in a time-invariant manner and has difficulty separating fixed effects from the forces that shape pathways and select cities into different growth experiences. An ideal way out of this dilemma would be to be able to estimate the sources for growth over a wide panel of cities for different time-periods. High-quality, sufficiently disaggregated, data on the nature of specialisation, human capital and institutions would be required, as well as those for a wide set of controls. Once the structural determinants for different periods are estimated and compared, then a further stage of research would consist in estimating them for individual city-regions. Finally, urban research would benefit from assembling rigorous data on events/shocks, not simply on panels of fixed effects, structural determinants or controls. For example, having data on the establishment and lock-in of localisations, or on other time-dependent shocks such as technologies that shock trade costs of existing sectors, or of changes in institutions that alter comparative advantages, and then estimating their effects jointly with structural determinants over different time-periods would get us much closer to measuring sequential (possibly path-determining) forces and structural determinants in a single model.

This would represent a significant departure from most of the research that is reported today, in that it would join approaches that are radically separated in academia: large-area quantitative analysis of structural determinants as favoured by economists and

large-area comparative development, as in the methodological tradition of Barrington Moore (1966), or Charles Tilly (1984), with the substantive contemporary concerns of comparative growth economics and the 'on the ground' sensibilities of urban geographers. Although 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 in understanding the complex problem of differential urban growth and the major transformation of urban systems.

Notes

1. I am referring to the notion, from international growth economics, that there will be strong tendencies to income convergence for countries with similar technological, legal and educational endowments, but weak tendencies among countries with highly unequal endowments, based on the notion that factor complementarities are strong within structural clubs, so that mobility smoothes out incomes, but weak between different clubs, so that it does not help much (Baumol and Wolff, 1988).
2. The recent literature on convergence and divergence of urban incomes includes, notably, Drennan *et al.*, 1996.
3. This notion does not figure prominently in general equilibrium approaches to economic geography. In those, movements in factor stocks and factor prices, combined with the absence of significant agglomeration economies, will eliminate any possible monopolistic competition effects over space (Anas *et al.*, 1998).
4. There is a different debate about whether it is better for a city to be specialised or diversified, but it should not be confused with the problem we are discussing here (Duranton and Puga, 2000).
5. This basic split has been identified, and interpreted quite differently, in Scott (2008).
6. Per capita income in the five-county Los Angeles metro area was 94 per cent that of the ten-county San Francisco metro area in 1970, but it had declined to 68 per cent by 2005.

7. Quality-of-life indices, used to instrument for 'amenities', include those of Mercer, the *Reader's Digest*, and *The Economist*. All rank the San Francisco Bay Area far higher in quality of life than the Los Angeles area.
8. In this model, as well as a large accompanying literature, the assumption is that regulation determines the elasticity of expansion of the housing base and hence of the labour force (Glaeser and Gyourko, 2002). However, an alternative (van Nieuwerburgh and Weill, 2009) assigns housing price differences to productivity differences, reflected in wage dispersion, between metropolitan areas. This would mean that the determinants of productivity would ultimately be of greatest interest, calls into question the three-equilibrium picture and leaves open the relationship of income growth to population growth.
9. In the US, the principal index used is the Wharton Index.
10. Heterogeneity could be simply due to the fact that the more people there are, the higher probability that their preferences will be dispersed; or because higher population raises the probability that it will involve ethnic, religious or linguistic fragmentation, which often corresponds to different preferences. Income inequality could also be a source of divergent preferences.
11. It is important to note that at the international scale, there is now a body of (imperfect) wide-area empirical testing of different hypotheses about how institutions affect growth and performance (Acemoglu *et al.*, 2004; Prezeworski *et al.*, 2000). At the inter-regional scale, however, while there is an enormous literature, it is almost entirely qualitative or case-study based; we lack any systematic evidence that institutions, politics and governance at the regional level actually matter to regional economic performance in the medium to long run.
12. Historical research using this approach can be found in Padgett and Ansell, 1993.
13. This is the opposite of the problem taken on by Davis and Weinstein (2002): they show the persistence of places after shocks as an example of path dependencies; we suggest that there are shocks that can create new places, but then lock in these shocks to make them durably important.
14. Another example of institutional problem-solving that shapes and shifts the geography of an industry would be the finance sector in Venice–Amsterdam–London–New York from the 16th century onwards.
15. Some attempts do exist in the literature. Most recently, Abu-Lughod (1999) attempts a comparison of development in New York, Chicago and Los Angeles, although her approach does not use a model with strictly common and comparable structural elements to analyse economic development.

Acknowledgements

This research was supported by a grant from the John Randolph Haynes and Dora Haynes Foundation. The author thanks four anonymous referees, as well as participants in seminars at Fudan University, Sciences Po, the National Taiwan University, the University of Utrecht, the University of Amsterdam, the Corporación Andina de Fomento (Lima) and the Instituto Fernando Henrique Cardoso (Rio de Janeiro) for comments on previous drafts.

References

- Abu-Lughod, J. (1999) *New York, Chicago, Los Angeles: America's Global Cities*. Minneapolis, MN: University of Minnesota Press.
- Acemoglu, D., Johnson, S. H. and Robinson, J. A. (2004) *Institutions as the fundamental cause of long-run growth*. Working Paper No. 10481, NBER, Cambridge, MA (www.nber.org/working-papers).
- Aghion, P., Alesina, A. F. and Trebbi, F. (2005) *Choosing electoral rules: theory and evidence from US cities*. Working Paper No. 11236, NBER, Cambridge, MA (www.nber.org/working-papers).
- Alesina, A. and Spolaore, E. (2003) *The Size of Nations*. Cambridge, MA: MIT Press.
- Amsden, A. H. (1989) *Asia's Next Giant: South Korea and Lake Industrialization*. Oxford: Oxford University Press.
- Anas, A., Arnott, R. and Small, K. (1998) Urban spatial structure, *Journal of Economic Literature*, 36, pp. 1426–1464.

- Barro, R. (1996) *The determinants of economic growth: a cross-country empirical study*. Working Paper No. 5698, NBER, Cambridge, MA (www.nber.org/working_papers).
- Baumol, W. J. and Wolff, E. N. (1988) Productivity growth, convergence and welfare: reply, *American Economic Review*, 78(5), pp. 1155–1159.
- Becattini, G. (1990) The Marshallian industrial district as a socio-economic notion, in: F. Pyke and W. Sengenber (Eds) *Industrial Districts and Inter-firm Cooperation in Italy*, pp. 37–51. Geneva: International Labour Office.
- Boschma, R. A. and Kloosterman, R. C. (Eds) (2005) *Learning from Clusters: A Critical Assessment from an Economic-geographical Perspective*. Frankfurt: Springer.
- Boschma, R. A. and Martin, R. (2007) Constructing an evolutionary economic geography, *Journal of Economic Geography*, 7(5), pp. 537–548.
- Carruthers, J. (2002) Growth at the fringe: the influence of political fragmentation in United States metropolitan areas, *Papers in Regional Science*, 82, pp. 475–499.
- Cheshire, P. and Sheppard, S. (1995) On the price of land and the value of amenities, *Economica*, 62, pp. 247–267.
- Combes, P. P. and Duranton, G. (2006) Labour pooling, labour poaching, and spatial clustering, *Regional Science and Urban Economics*, 36(1), pp. 1–28.
- Cox, K. R. (1993) The local and the global in the new urban politics: a critical view, *Environment and Planning D*, 11(4), pp. 433–448.
- Cusumano, M. J. (1985) *The Japanese Automobile Industry: Technology and Management at Nissan and Toyota*. Cambridge, MA: Harvard University Press.
- Dahl, R. (1961/2005) *Who Governs? Democracy and Power in an American City*, 2nd edn. New Haven, CT: Yale University Press.
- Davis, R. D. and Weinstein, D. E. (2002) Bones, bombs, and break points: the geography of economic activity, *American Economic Review*, 92(5), pp. 1269–1289.
- Djankov, S., Glaeser, E., LaPorta, R. *et al.* (2003) The new comparative economics, *Journal of Comparative Economics*, 31(4), pp. 595–619.
- Drennan, M. (2002) *The Information Economy and American Cities*. Baltimore, MD: Johns Hopkins University Press.
- Drennan, M., Tobier, E. and Lewis, J. (1996) The interruption of income convergence and income growth in large cities in the 1980s, *Urban Studies*, 33(1), pp. 63–82.
- Duranton, G. and Puga, D. (2000) Diversity and specialisation in cities: why, where and when does it matter?, *Urban Studies*, 37(3), pp. 533–555.
- Duranton, G. and Puga, D. (2001) Nursery cities: urban diversity, process innovation, and the life cycle of products, *American Economic Review*, 91(5), pp. 1454–1477.
- Duranton, G. and Puga, D. (2004) Micro-foundations of urban agglomeration economies, in: J. V. Henderson and J.-F. Thisse (Eds) *Handbook of Regional and Urban Economics, Vol. 4, Cities and Geography*, pp. 2064–2117. Amsterdam: Elsevier.
- Ellison, D. J., Clark, R. B., Fujimoto, T. and Hyun, Y. S. (2002) *Product development performance in the automobile industry: 1990s updates*. Working paper, MIT, Cambridge, MA.
- Erie, S. P. (1990) *Rainbow's End: Irish-Americans and the Dilemma of Urban Machine Politics, 1845–1985*. Berkeley, CA: University of California Press.
- Florida, R. (2002) *The Rise of the Creative Class*. New York: Basic Books.
- Fuchs, G. and Shapira, P. (Eds) (2005) *Rethinking Regional Innovation and Change: Path Dependency or Regional Breakthrough*. Frankfurt: Springer.
- Fujita, M. and Thisse, J. (2002) *Economics of Agglomeration: Cities, Industrial Location and Regional Growth*. Cambridge: Cambridge University Press.
- Galbraith, J. and Hale, G. (2004) *Income distribution and the information technology bubble*. Working Paper No. 27, University of Texas Inequality Project, Austin, TX.
- Glaeser, E. L. (2007) *The economic approach to cities*. Department of Economics, Harvard University, Cambridge, MA (<http://www.economics.harvard.edu/faculty/glaeser/files/kayden4.pdf>).
- Glaeser, E. L. and Gyourko, J. (2002) *The impact of zoning on housing affordability*. Harvard Institute of Economic Research, Cambridge, MA.
- Glaeser, E. and Shapiro, J. (2003) Urban growth in the 1990s: is city living back?, *Journal of Regional Science*, 43(1), pp. 139–165.

- Glaeser, E., Kolko, J. and Saiz, A. (2001) Consumer city, *Journal of Economic Geography*, 1(1), pp. 27–50.
- Glaeser, E., Scheinkman, J. and Schleifer, A. (1995) Economic growth in a cross section of cities, *Journal of Monetary Economics*, 36(1), pp. 117–143.
- Glaeser, E., Kellal, H. D., Scheinkman, J. A. and Schleifer, A. (1992) Growth in cities, *Journal of Political Economy*, 100(6), pp. 1126–1152.
- Glaeser, E., LaPorta, R., Lopez-de-Silanes, F. and Schleifer, A. (2004) Do institutions cause growth?, *Journal of Economic Growth*, 9(3), pp. 271–303.
- Granovetter, M. (1995) *Getting a Job: A Study of Job Contacts and Careers*. Chicago, IL: University of Chicago Press.
- Granovetter, M. (2005) The impact of social structure on economic outcomes, *Journal of Economic Literature*, 19(1), pp. 33–50.
- Grossman, G. M. and Helpman, E. (1991) Quality ladders and the theory of growth, *Review of Economic Studies*, 58(1), pp. 43–61.
- Hall, P. and Soskice, D. (Eds) (2001) *Varieties of Capitalism*. Oxford: Oxford University Press.
- Hausman, R. and Rodrik, D. (2003) Self-discovery in a development strategy for El Salvador, *Economia: Journal of the Latin American and Caribbean Economic Association*, 6(1), pp. 43–102.
- Hilber, C. A. J. and Mayer, C. J. (2004) *Why do households without children support local public schools?* Working Paper No. 10804, NBER, Cambridge, MA, October (www.nber.org/papers/w10804).
- Hodgson, G. (1993) *Economics and Evolution*. Ann Arbor, MI: University of Michigan Press.
- Jaher, F. C. (1984) *The Urban Establishment: Upper Strata in Boston, New York, Charleston, Chicago and Los Angeles*. Urbana, IL: University of Illinois Press.
- Jayet, H. (1983) Chômeur plus souvent en région urbaine, plus longtemps en région rurale [Labor turnover is greater in urban areas, but unemployment more durable in rural areas], *Economie et Statistique*, 153, pp. 47–57.
- Jones, B. D. and Baumgartner, F. (2005) *The Politics of Attention: How Government Prioritizes Problems*. Chicago, IL: University of Chicago Press.
- Kennedy, M. (Ed.) (2000) *Understanding Silicon Valley*. Stanford, CA: Stanford University Press.
- Kenyon, D. (1997) Theories of interjurisdictional competition, *New England Economic Review*, March/April, pp. 14–28.
- Kenyon, D. and Kincaid, J. (Eds) (1991) *Competition among States and Local Governments*. Washington, DC: Urban Institute Press.
- Krugman, P. (1991a) Increasing returns and economic geography, *Journal of Political Economy*, 99, pp. 484–499.
- Krugman, P. (1991b) History versus expectations, *Quarterly Journal of Economics*, 106(2), pp. 651–667.
- Krugman, P. and Obstfeld, M. (1991) *International Economics: Theory and Policy*. New York: Harper Collins.
- Logan, J. and Molotch, H. (1987) *Urban Fortunes*. Berkeley, CA: University of California Press.
- Lucas, R. (1988) On the mechanics of economic development, *Journal of Monetary Economics*, 22, pp. 3–42.
- Mokyr, J. (1990) *The Lever of Riches: Technological Creativity and Economic Progress*. New York: Oxford University Press.
- Molotch, H. (1976) The city as a growth machine, *American Sociological Review*, 82(2), pp. 309–332.
- Moore, B. (1966) *Social Origins of Dictatorship and Democracy*. Boston, MA: Beacon Press.
- Muth, R. F. (1971) Migration: chicken or egg?, *Southern Economic Journal*, 37, pp. 295–306.
- Niewerburgh, S. van and Weill, P.-O. (2009) *Why has house price dispersion gone up?* Department of Economics, University of California, Los Angeles (<http://www.econ.ucla.edu/poweill/houseprice.pdf>).
- North, D. (2005) *Understanding the Process of Economic Change*. Princeton, NJ: Princeton University Press.
- Norton, R. D. and Rees, J. (1979) The product cycle and the spatial decentralization of American manufacturing, *Regional Studies*, 13(2), pp. 141–151.
- Oates, W. (1969) Local spending on property values, *Journal of Political Economy*, 6(November/December), pp. 957–971.
- Padgett, J. F. and Ansell, C. K. (1993) Robust action and the rise of the Medici, 1400–1434, *American Journal of Sociology*, 98(6), pp. 1259–1319.
- Padgett, J. F. and McLean, P. (2006) Organizational invention and elite transformation: the birth of the partnership system in renaissance Florence, *American Journal of Sociology*, 111(5), pp. 1463–1568.

- Persson, T. and Tabellini, G. (2006) *Democracy and development: the devil is in the details*. Working Paper No. 11993, NBER, Cambridge, MA (www.nber.org/papers/wp11993).
- Peterson, P. (1981) *City Limits*. Chicago, IL: University of Chicago Press.
- Powell, W., Packer, K. and Whittington, K. (forthcoming) Organizational and institutional genesis: the emergence of high-tech clusters in the life sciences, in: J. Padgett and W. Powell (Eds) *The Emergence of Organization and Markets*, ch. 14. (in preparation).
- Powell, W. W., White, D. R., Koput, K. W. and Owen-Smith, J. (2005) Network dynamics and field evolution: the growth of inter-organizational collaboration in the life sciences, *American Journal of Sociology*, 110(4), pp. 1132–1205.
- Pritchett, L. (1997) Divergence, big time, *Journal of Economic Perspectives*, 11(3), pp. 3–17.
- Przeworski, A., Alvarez, M., Cheibub, J. A. and Limongi, F. (2000) *Democracy and Development: Political Institutions and Well Being*. Cambridge: Cambridge University Press.
- Putnam, R. D. (2000) *Bowling Alone: The Collapse and Revival of American Community*. New York: Touchstone.
- Rigby, D. L. and Essletzbichler, J. (1997) Evolution, process variety and regional trajectories of technological change in US manufacturing, *Economic Geography*, 73, pp. 269–284.
- Rodrik, D. (2007) *One Economics, Many Recipes: Globalization, Institutions and Economic Growth*. Princeton, NJ: Princeton University Press.
- Rodrik, D., Subramanian, A. and Trebbi, F. (2004) Institutions rule: the primacy of institutions over geography and integration in economic development, *Journal of Economic Growth*, 9(2), pp. 131–165.
- Rose-Ackerman, S. (1983) Beyond Tiebout: modeling the political economy of local government, in: G. Zodrow (Ed.) *Local Provision of Public Services: The Tiebout Model after Twenty-five Years*, pp. 55–83. New York: Academic Press.
- Rosen, S. (1983) Specialization and human capital, *Journal of Labor Economics*, 1(1), pp. 43–49.
- Rosenthal, S. R. and Strange, W. C. (2001) The determinants of agglomeration, *Journal of Urban Economics*, 53, pp. 300–320.
- Saks, R. E. (2007) Job creation and housing construction: constraints on metropolitan area growth, *Journal of Urban Economics*, 64, pp. 178–195.
- Sala-i-Martin, X. (2002) *15 years of new growth economics: what have we learnt?* Discussion Paper No. 0102-47, Department of Economics, Columbia University.
- Saxenian, A. L. (1994) *Regional Advantage: Culture and Competition in Silicon Valley*. Cambridge, MA: Harvard University Press.
- Saxenian, A. L. (2000) The origins and dynamics of production networks in Silicon Valley, in: R. Swedberg (Ed.) *Entrepreneurship: The Social Science View*, pp. 384–403. Oxford: Oxford University Press.
- Scott, A. J. (1993) *Technopolis: High Technology Industry and Regional Development in Southern California*. Berkeley, CA: University of California Press.
- Scott, A. J. (2008) Production and work in the American metropolis: a macroscopic approach, *Annals of Regional Science*, 42, pp. 785–805.
- Scott, A. J. and Storper, M. (1987) High technology industry and regional development: a theoretical critique and reconstruction, *International Social Science Journal*, 112(May), pp. 215–232.
- Stiglitz, J. (1983) The theory of local public goods twenty-five years after Tiebout, in: G. Zodrow (Ed.) *Local Provision of Public Services: The Tiebout Model after Twenty-five Years*, pp. 17–53. New York: Academic Press.
- Storper, M. (2005) Society, community and economic development, *Studies in Comparative International Development*, 39(4), pp. 30–57.
- Sveikauskas, L. (1975) The productivity of cities, *Quarterly Journal of Economics*, 89(3), pp. 393–413.
- Tilly, C. (1984) *Big Structures, Large Processes, Huge Comparisons*. New York: Russell Sage Foundation.
- Trefler, D. (1993) International factor price differences: Leontief was right!, *Journal of Political Economy*, 101(6), pp. 961–987.
- Turner, F. (2006) *From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Catalog, and the Rise of Digital Utopianism*. Chicago, IL: University of Chicago Press.