Agglomerated, Built, and Clustered: The ABCs of Infrastructure, Spillover, and Innovation

Adam Koling, Daniel Armanios, Agni Orfanoudaki (Saïd Business School, Univ. of Oxford)

Research Problem Statement

Density, urbanization, and geographic co-location are implicitly or explicitly prioritized in organizational decisions, and cited in literature, as key drivers of local and regional economic productivity and innovation [1]. Multiple proposed mechanisms may explain the outsized efficiency and performance achieved in cities, from tangible spatial connectivity to reduced friction in the flow of knowledge and ideas [2]. It follows intuitively that physical environments and knowledge flows work together, with simultaneous movements of ideas and people as symbiotic forces in regional economies. We know relatively little, however, about which economic conditions and types of built environments are most suited to one another, and whether there are optimal combinations to activate knowledge spillovers and ensuing innovations.

In this working paper, we consider the role of urban built environments, in particular transport, land use, and broadband systems, and how they interact with two contrasting knowledge flow theories: Jacobs and Marshall, Arrow, Romer (MAR). We pose three research questions: first, whether industry specialization and market competition influence innovation outcomes at the county level in the United States, evaluating evidence for the Jacobs and MAR theories across the United States. Second, we test whether Jacobs and MAR economic conditions matter more in regions characterized by one infrastructure genre over another. Finally, we deploy the same techniques to explore whether regions with higher digital connectivity tend to favor Jacobs or MAR spillovers.

Knowledge spillovers, where the movement and exchange of ideas from one organization stimulate innovation or productivity in one or more neighboring organizations, are difficult to measure directly. Krugman (1991) cautions that unlike material agglomeration mechanisms such as labor pools or supply chains, “knowledge flows, by contrast, are invisible; they leave no paper trail.” [3] Observing such spillovers in real time has been, and the primary approach taken in most literature is to evaluate various metrics of innovative output and economic performance and to infer knowledge flows based on outcomes. Agglomeration theories assume information flows cause knowledge spillovers, yet these are precisely the processes researchers cannot easily measure, and empirical analyses to date have shown mixed results [4].

Three core spillover theories appear throughout agglomeration literature, and, more recently, in engineering and infrastructure literature exploring a nexus with the built environment and innovation and entrepreneurship. Each is named after its proponents: (a) Jacobs (1969), (b) MAR (1890, 1962, and 1986, respectively), and (c) Porter (1990), and proposes different combined

conditions of market competition and industry specialization for knowledge. We focus only on the first two Jacobs and MAR, that contrast more directly with each other [5]. Jacobs spillovers occur in regions with high competition (many firms) and high diversification (many industries), and MAR spillovers occur in the opposite local context: monopoly or oligopoly (fewer firms) and higher specialization (many industries).

Research Methodology and Approach

If Jacobs spillover conditions enable more economic activity, then regions displaying both higher market competition and greater industry diversification would be expected to display stronger productivity and innovation outcomes. Conversely, if more pronounced MAR conditions drive innovation, then we would expect these outcomes in regions with lower market competition and less industry diversification. Our first hypotheses directly pertain to MAR and Jacobs spillover predictions, at the county level, across the United States, using a comprehensive panel dataset from County Business Patterns (CBP). We test this by a series of regressions, in which our overall dataset is split by percentile into MAR and Jacobs quadrants, and further subset into *strong* MAR and Jacobs quadrants whose units fall in the top quartile of specialization, concentration, diversification, and/or competition, respectively.

Jacobs spillovers are predicated in part on serendipitous interactions, which should increase with more diverse social and organizational life, likely under conditions of greater diversification across more firms [6]. Therefore, multi-modal transport networks, against a backdrop of mixed land use and density, would deliver more channels for serendipitous interaction. While MAR spillover literature does not refer as explicitly to built environments as Jacobs does, the underlying mechanisms emphasize, we argue that MAR spillovers are more strongly associated with agglomeration benefits achieved in more secluded or divided urban areas. To strike this balance between isolation (to prevent leakage) and proximity (to allow for directed and frequent interaction), we expect stronger innovation and productivity outcomes in more suburban or auto-dominated regions, where firms and people can isolate on office campuses yet commute and travel at will [7]. Our second set of hypotheses evaluate the same outcomes while introducing variables that characterize the transport and land use context, aggregating from various spatial datasets.

Our third set of hypotheses extends the same methodology as the second set, but examining digital connectivity, broadband internet access and uptake, rather than physical and transport infrastructure [8]. We expect access to and consumption of novel information via high-speed web access to complement MAR spillover mechanisms, because in-person interactions can be more deliberate and fruitful, and rely less on spontaneity and serendipity than Jacobs spillovers [9]. Therefore, we test overall regional broadband access and recorded use.

We deploy both multiple linear regression and quasi-Poisson estimation techniques to test these hypotheses cross-sectionally. At the time of this extended abstract, we have evaluated them for four outcome variables: (a) new hires, (b) new firm creation, (c) new hires in new firms, and (d) patents all at the United States county level between 2017 and 2021. Our infrastructure variables include (a-c) auto, multimodal, and pedestrian link densities, (d-e) auto and nonauto intersection densities, (f-g) auto and non-auto access indices, (h) public transport regional centrality indices, and (i) walkability indices, aggregated manually by county from GIS files downloaded from the Smart Location Database of the United States Environmental Protection Agency (EPA). Digital connectivity is measured by (j) federal availability records and (k) publicly-available Microsoft data for usage. Where necessary, block group transport and infrastructure metrics are aggregated to county level by an average weighted by the quantity of housing units. All infrastructure and broadband variables are brought into multiple linear regression and quasi-Poisson models as interaction effects.

Preliminary Results and Implications

While we have yet to complete all model variations, at the time of this extended abstract, our initial results indicate consistent positive and statistically significant interaction effects of spillover conditions and pedestrian infrastructure, with a higher proportion of mixed or null results for auto and multimodal infrastructure. While there are statistically significant results, particularly for walkability and pedestrian infrastructure, there is no immediate pattern of MAR and Jacobs quadrants influencing outsized performance in the presence of a certain infrastructure type. These results suggest that knowledge flows may first be unlocked at the microgeographic level, and that the theorized relationships between regional infrastructure types and spillover mechanisms may not account sufficiently for much smaller-scale human connectivity. Broadband access and usage do not exhibit clear patterns for firm or job creation variables, but evidently favor MAR quadrants for patents. With no infrastructure variables, there is a persistent and statistically significant difference between outcomes in MAR and Jacobs quadrants, with employment and firm creation outcomes favoring Jacobs and patents favoring MAR. Consequently, our first hypothesis is supported, and our second and third partially supported.

Our results also appear to be robust to model type (multiple linear regression vs. quasi-Poisson), robustness checks with outlying counties excluded, as well as inclusion of each infrastructure variable as additive controls vs. interactions in our multiple linear regression models. Subsequent steps involve testing these relationships ourselves using the Google Distance Matrix API to compare pedestrian, automobile, and public transport travel times between random points within the spatial units in question, introducing several additional outcome variables such as startup formation, and expanding on the National Walkability Index by including one or more different aggregated walk and transit scores for each unit to our infrastructure variables.

These findings deliver relevant information not only to scholars of cities, agglomeration economies, and physical infrastructure, but also to the engineering and planning professions writ large. Land use, transport, and urban development are not simply characterized by density and proximity, but also by market conditions and industry types present. Ultimately, the performance of these industries (by a variety of outcome metrics) hinges on correct measurement and interpretation of their relationships with human and built environments.

References

[1] Ellison, Glenn, and Edward L. Glaeser. "The geographic concentration of industry: does natural advantage explain agglomeration?." American Economic Review 89, no. 2 (1999): 311-316.
[2] Marshall, Alfred. "1920. Principles of economics." London: Mac-Millan (1890): 1-627.
[3] Krugman, Paul. Geography and trade. MIT press, 1992.
[4] Beaudry, Catherine, and Andrea Schiffauerova. "Who's right, Marshall or Jacobs? The localization versus urbanization debate." Research policy 38, no. 2 (2009): 318-337.
[5] Nica, Mihai. "Small business clusters in Oklahoma: Mar or Jacobs effects." Regional and Sectoral Economic Studies 10, no. 2 (2010): 5-19.
[6] Desrochers, Pierre, and Samuli Leppälä. "Opening up the 'Jacobs Spillovers' black box: local diversity, creativity and the processes underlying new combinations." Journal of Economic Geography 11, no. 5 (2011): 843-863.
[7] Carlino, Gerald A. "Knowledge spillovers: cities' role in the new economy." Business Review Q 4, no. 1 (2001): 17-24.
[8] Ritsch, Nicola, and Daniel Erian Armanios. "Using broadband infrastructure as a social sensor to detect inequities in unemployment during the COVID-19 pandemic." Scientific Reports 13, no. 1 (2023): 22031.
ix Cariolle, Joël, and Maëlan Le Goff. "Spatial internet spillovers in manufacturing." The Journal of Development Studies 59, no. 8 (2023): 1163-1186.