Immigrant Diversity and Economic Development in Cities: A Critical Review

Thomas Kemeny (LSE)

November 2013
This work is part of the research programme of the independent UK Spatial Economics Research Centre funded by a grant from the Economic and Social Research Council (ESRC), Department for Business, Innovation & Skills (BIS) and the Welsh Government. The support of the funders is acknowledged. The views expressed are those of the authors and do not represent the views of the funders.

© T. Kemeny, submitted 2013
Immigrant Diversity and Economic Development in Cities: A Critical Review

Thomas Kemeny*

November 2013

* Department of Geography and Environment, LSE

Acknowledgements
This research was supported by funding from the Regional Studies Association’s Early Career Grant.
Abstract
This paper reviews a growing literature investigating how ‘immigrant’ diversity relates to urban economic performance. As distinct from the labor-supply focus of much of the economics of immigration, this paper reviews work that examines how growing heterogeneity in the composition of the workforce may beneficially or harmfully affect the production of goods, services and ideas, especially in regional economies. Taking stock of the existing literature, the paper argues that the low-hanging fruit in this field has now been picked, and lays out a set of open issues that need to be taken up in future research in order to fulfil the promise of this work.

Keywords: diversity, immigration, cities, regional economic performance
JEL classification: O4, O15, O18, O31, R0, J28, J31
1 Introduction

Immigration is making many countries increasingly diverse. This growth in diversity is rooted in changes in the size of aggregate immigrant flows, as well as changes in their composition. In the United States, for instance, there were over four times as many foreign-born residents in 2011 as there were in 1970, and the share of immigrants recently surpassed 13 percent, approaching the nation’s historical peak, occurring almost one hundred years ago. Although the U.S. remains the primary global immigrant destination, absorbing one in five international migrants, rising immigration is by no means exclusively an American phenomenon. Across all OECD countries, foreign-born workers now account for about one-tenth of the workforce (Alesina et al., 2013). Worldwide, between 1960 and 2000, the number of international migrants more than doubled, with a disproportionate share headed to advanced economies (Özden et al., 2011). Over and above the absolute growth in migration flows, the range of countries from which migrants hail has broadened. Migration flows are made up of a greater variety of intra-regional, as well as South-North, North-North and South-South moves (ibid). Growing international migration has combined with this growing heterogeneity of source countries to considerably expand the cultural diversity of recipient nations. This immigrant-induced diversity is not evenly distributed – it is especially concentrated in metropolitan areas. In ‘global’ cities like New York, Los Angeles, London, and Hong Kong, immigrants now make up more than one-third of the populace.

Political, social and economic effects of immigration have been much debated among academics (Borjas, 1994; Freeman, 1995; Alba and Nee, 1997), and immigration is a ubiquitous hot-button topic in the popular media in the U.S. and Europe. Scholarly economic perspectives have focused on the two extreme ends of the labor market: the impact of low-skill immigrants on low-skill natives (Card, 2005; Cortes, 2008), and the much less controversial outcomes associated with the immigration of highly-skilled workers (Borjas, 2005; Saxenian, 2007). At both ends, however, researchers have mostly approached the issue of immigration’s impact as one that rests upon changes wrought to aggregate labor supply. Less well understood are the economic effects of interactions among a diverse populace, and specifically how these interactions might affect the production of goods, services and ideas. However, over the past decade, economic geographers, urban economists, and development economists have taken up this subject, seeking to better
understand how regional and national economies might perform differently when composed of heterogeneous workers. This heterogeneity has been often (though not exclusively) defined in terms of immigration.

The purpose of this article is to review this growing literature, focusing especially on theory and quantitative empirical work seeking to understand the economic impacts of immigrant diversity at the metropolitan scale. The studies to be reviewed are motivated by theory in which the presence of immigrant-diverse individuals could simultaneously improve economic outcomes by bringing together different perspectives and heuristics, and reduce performance by making co-operation more costly. Based on variants of this theory, the literature examining the metropolitan and national effects of ‘birthplace’ or ‘immigrant’ diversity has found suggestive evidence of a positive relationship between diversity and three outcomes: productivity, innovation and entrepreneurship.

Interest in these relationships has grown over nearly a decade; they have recently been the subject of special issues and dedicated conferences in economic geography and regional science; and there is now a substantial body of scholarship, with emergent norms and practices, that is worthy of careful consideration. It is an appropriate moment to investigate how existing research has approached the study of immigrant diversity and regional economic development, and what challenges remain to be addressed.

This review finds several significant conceptual and methodological issues that currently stand in the way of making confident statements about economic benefits and costs arising from immigrant heterogeneity in urban and regional contexts. First, extant work has not sufficiently dealt with inter-urban sorting dynamics that could pull unobservably higher skilled workers to cities that are also immigrant-diverse. Second, studies have largely relied upon problematic instruments in order to deal with potential reverse causation in the links between diversity and performance. Third, more work is needed that explores how urban immigrant diversity may have effects that vary across different labor market segments. Fourth, studies should also explore how the hypothesized double-edged relationships between diversity and economic performance may be moderated by institutions and other social and spatial forces. Fifth, economic geographers must also pay closer heed to diversity-focused research on organizations, as well as in developing countries.

---

1 This article will use the terms ‘regional,’ ‘metropolitan,’ ‘city,’ and ‘urban’ more or less interchangeably to refer to functionally-linked economic regions, as distinct from, say, municipalities demarcated for administrative purposes.
nally, researchers must better understand the mechanisms driving any potential outcomes, and be able to more precisely measure the independent variable in question. Is urban immigrant heterogeneity a proxy for an omitted variable, and if so, what is that force and how does it operate upon productivity, innovation and entrepreneurship. By addressing these issues, the field can enhance causal influence; produce better estimates of the relationships in question; and open up new areas of inquiry that help us understand the broader impacts of immigrant heterogeneity in urban (and other) economies.

2 Diversity and Economic Performance: A Theoretical Framework

Studies of the economic impacts of diversity have been motivated by a large, diverse and well-established body of theoretical and empirical research, whose contributors span such fields as psychology, organizational studies, artificial intelligence and economics. The chief aim of this literature has been to understand whether and how formal work teams composed of heterogeneous agents may be more effective than those that are homogenous.

From this workgroup and organizational diversity literature comes the central, double-edged theoretical predictions linking heterogeneity and performance. One side of this argument hypothesizes that, by enlarging the pool of available perspectives and heuristics, groups that are diverse should outperform those that are homogenous (Nisbett et al., 1980; Clearwater et al., 1991; Thomas and Ely, 1996; Hong and Page, 2001). Elucidating the mechanism behind this effect, Hong and Page (2004) consider that individuals with identity diversity, defined as those with particular demographic, geographic, ethnic, or cultural backgrounds, are also likely to be distinctive in terms of their functional diversity, meaning the ways they perceive and solve problems. This hypothesized link between identity and cognitive function forms the basis for advantages that could emerge as diverse individuals interact. Assuming that a given challenge can be overcome in multiple ways, a group that has access to a larger number of perspectives on the problem, as well as approaches to its resolution, ought to adopt a more effective solution. This benefit arises for two reasons. First, diverse agents in cooperation will jointly map out a larger proportion of the potential solutions available in the total problem space. Second, functionally-diverse agents can cross-pollinate,
yielding novel solutions – innovations – that are not directly a function of any singular set of perspectives and heuristics (Aiken and Hage, 1971). In mathematical models of the underlying mechanism, researchers find that diverse groups of problem solvers can even outperform teams that consist of agents with superior abilities, since best-performing individual agents tend to experience a convergence in functional diversity (Huberman, 1990; Hong and Page, 2004). Empirical work exploring the economic implications of identity diversity have mostly focused on interactions between diverse individuals in the workplace. Exploring a wide variety of forms of identity diversity, this literature has found a wealth of evidence in support of positive, if modest economic effects (Hoffman and Maier, 1961; Bantel and Jackson, 1989; Lazear, 1999; Herring, 2009; Joshi and Roh, 2009).

The counter-hypothesis is rooted in psychology’s ‘social identity theory,’ which predicts that the presence of diverse individuals within teams will stimulate the formation of informal intra-team groups, leading to a situation in which in-group members will be favored and trusted more than members of out-groups (Tajfel, 1974; Abrams and Hogg, 1990). This internal fractionalization can hinder cooperation and promote rent-seeking behavior, thus reducing productivity (Byrne, 1971; Turner et al., 1987; Chatman and Flynn, 2001; Van Knippenberg and Schippers, 2007; Harrison and Klein, 2007). Various evidence exists to support this hypothesis. For instance, research indicates that team members who share few commonalities find it hard to integrate and effectively communicate (Richard et al., 2002; Ancona and Caldwell, 1992). Diverse work groups have also been associated with reduced co-operation (Bandiera et al., 2005), as well as higher levels of dissatisfaction and turnover (O’Reilly et al., 1989).

While these two conceptions of the effects of diversity are oppositional, their underlying mechanisms are not incompatible. The positive view is that diversity introduces production complementarities that, according to the organizational literature at least, extend to workgroup members and thus to the organization in which that team operates. The negative view observes that diversity tends to obstruct collaboration and cooperation, inhibiting team performance. Yet both of these mechanisms can operate simultaneously. Diversity’s benefits may be real but latent, depending on particular social arrangements for their realization. Rather than presuming diversity to be strictly good or bad for productivity (Stahl et al., 2009), it may make more sense to think of
diversity’s effects as being contingent upon transaction costs. When the cost of transacting across either identity or functional forms of diversity is sufficiently low, the benefits of diversity may shine through. When these transaction costs are high, they may choke off any benefits latent in collective heterogeneity. Under this view, the outcome depends on the costs of interacting in a specific context.

Researchers studying regional and national economies have transposed this model of diversity’s economic effects from the workgroup to the city, on the basis that the public-good characteristics of diversity could plausibly spill beyond teams and organizations. Economies require interaction and coordination across work teams and between atomized firms (North, 1990; Storper, 1997), hence highly fractionalized locations could generate negative externalities that might hinder development. This is precisely the line of argument theorized and confirmed at the national scale in recent work by development economists (Alesina and Drazen, 1991; Easterly and Levine, 1997; Rodrik, 1999; Alesina and La Ferrara, 2005; Montalvo and Reynal-Querol, 2005). Researchers find that heterogeneity can also be a source of social conflict and economic underperformance at the urban scale. For instance, evidence suggests that U.S. cities and states fragmented by ethnicity or age spend less than more homogenous regions on productive public goods, such as roads, hospitals and schools (Poterba, 1997; Goldin and Katz, 1999; Alesina et al., 1999; Pennant, 2005).

It is not just the negative effects of heterogeneity that make sense beyond organizations; positive production externalities conceptualized at the workgroup level are also plausible at the regional scale. As Jane Jacobs (1969) famously observed, cities are engines of economic growth precisely because they enable and encourage people with different ideas to interact, resulting in new, economically-important knowledge. These interactions certainly do occur within organizations, but they also spill over beyond those confines (Glaeser et al., 1992; Saxenian, 1996; Feldman and Audretsch, 1999; Duranton and Puga, 2001; Ozgen et al., 2011a). The production of knowledge is best understood as a geographical phenomenon, as opposed to one occurring inside individual atomized firms (Audretsch and Feldman, 2004). And while knowledge spillovers extend past organizational boundaries, they attenuate across space (Rosenthal and Strange, 2008). To the extent that knowledge spillovers are rooted in cities, it is because urban regions are functionally-integrated economic units, structured by repeated face-to-face contact and shared conventions (Jaffe et al.,
Diversity can be considered as a particular form of human capital externality, and may operate in a manner similar to local education spillovers. Indeed, just as we know that there are rewards for workers who inhabit cities in which levels of skill and education are high (Lucas, 1988; Rauch, 1993; Moretti, 2004), workers in highly diverse cities may be rendered more productive and innovative than comparable workers inhabiting less diverse places.

It has not been lost on urban researchers that immigrant diversity may influence not only the goods, services and ideas that urban workers produce, but also their experience of everyday life in ways that may find economic expression. Immigrant diversity’s economic effects, in other words, need not be confined to the sphere of production; rather they may also play a role in consumption, worker satisfaction, and migration. Florida (2002; 2004), for instance, contends that correlates of diversity, ranging from the presence of an eclectic mix of ethnic restaurants to the presence of an urban populace with generally tolerant attitudes toward immigrants, can function as urban amenities. Cities that feature these diversity-induced qualities in abundance may disproportionately draw highly skilled workers whose preferences include a love of variety. And, in a somewhat similar vein, a large subset of modern urban economics is built on the idea that workers are willing to trade some part of their wages in exchange for the ability to consume location-specific features that raise their quality of life (Roback, 1982; Glaeser and Gottlieb, 2009).

Hence, just as diversity may impact worker productivity, it can also function as an amenity that (some) workers desire to consume, with implications for factor prices, in terms of the wages workers earn and the costs they face in the housing market.

It is these ideas around production and consumption that have motivated the growing empirical literature on the economic effects of immigrant diversity in cities. Existing work has focused on three outcomes: productivity, innovation and entrepreneurship. It is to this work that the paper now turns, beginning with how existing studies have operationalized the concept of diversity.

3 Measuring Diversity

Merriam-Webster’s dictionary defines *diversity* as “the condition of having or being composed of differing elements.” From a theoretical standpoint, all sorts of elements could be the basis for identity diversity. Yet, existing research on diversity in cities has focused particularly on diversity
defined through national origin.\textsuperscript{2} There are sensible reasons for doing so. Arguing for birthplace diversity over linguistic or racial fractionalization, Alesina et al. (2013) argue that “shaped by different education systems and social values, this type of diversity is more likely to result in production function complementarities than differences in skin color or language spoken at home” (p.6). Doubtless, individuals born in a particular country are profoundly shaped by their immersion in a distinctive institutional, social and cultural environment. These forces are not alone in acting upon individuals’ identity, but a plausible case can be made that they help shape their world view, and thus their ways of framing and solving problems.

Still, some clear lines can and should be drawn around this form of diversity, as against broader notions of ‘culture’ with which it has sometimes been conflated. Culture is an amorphous term, but we can be sure that national origin is insufficient to contain it. For one thing, many national economies contain different cultures; one would certainly want to distinguish between, for instance, francophone and anglophone cultures in Canada, or, holding language constant, between the distinct regional identities that emerge from Galicia, Andalucia and Madrid in Spain. As well, birthplace does not capture the phenomenon of second-generation immigrants, through which new cultural combinations are born.

This straightforwardness of diversity’s dictionary definition, and the relative ease of measuring one’s country of birth notwithstanding, immigrant diversity remains a latent concept, and thus not something that can be uniquely or precisely identified. Nonetheless, a measure of immigrant diversity should describe at least two characteristics of a distribution of individuals. First, it must capture the degree to which the foreign-born are present in a location. Second, it should describe the breadth of source countries from which those individuals originate.

The standard tool to measure categorical forms of diversity, whether of birthplaces, languages, occupations or other characteristics, is the Fractionalization index (for some examples, see Taylor and Hudson, 1972; Easterly and Levine, 1997; Knack and Keefer, 1997; Ottaviano and Peri, 2006;\textsuperscript{2}

\textsuperscript{2}This is true, to a lesser extent, in studies at the national level.
where $n$ is the proportion of residents in city $j$ who were born in country $r$; and $R$ is the number of different countries represented among residents of that city. The index will near zero as diversity decreases; its maximum value approaches one as heterogeneity increases. It measures the probability that two randomly-drawn individuals in a location were born in different countries. Insofar as it captures both the depth and breadth of immigrant groups in cities, the Fractionalization index has obvious appeal. There are, however, other choices in measuring immigrant and other forms of diversity in cities. Researchers have commonly used the simple fraction of foreign-born workers in the urban population as a measure of diversity. This has the obvious problem that it captures only immigration undifferentiated by source. Another indicator is the Shannon entropy index, commonly used to describe the diversity of species in ecology, and also widely used in the workgroup diversity literature. This index is calculated as follows:

$$Shannon_j = - \sum_{r=1}^{R} n_j^r \ln(n_j)$$  

Though results produced using Shannon and Fractionalization measures ought to be strongly correlated, they are not identical. The Shannon index will be more sensitive to departures from distributions in which the components are of unequal size, whereas the Fractionalization better represents diversity in situations where groups are of roughly similar size (Dawson, 2012; Taagepera and Ray, 1977). In other words, neither measure is better or worse at gauging diversity – each is best suited to a particular set of circumstances. In the context of birthplace diversity in cities or national economies, the Shannon index seems most likely to be useful, as regional and national economies tend to be composed of native-born residents that typically dominate in numerical terms, combined with a smaller subset of different immigrant populations of varying sizes. Some studies have also measured immigrant diversity using the simple share of foreign born in a location, as well as measures of polarisation.

---

3 Researchers examining diversity in the international context have also sought to capture the idea of polarization, using measures derived from Esteban and Ray (1994), and designed to capture the extent to which a distribution of ethnic groups approaches two evenly sized groups. However, the theoretical motivation for measuring polarisation is substantially different from that of diversity in the present context; the latter is fundamentally about possibilities for conflict and rent seeking, while there is little conceivable benefit that could arise from either more or less polarisation in an urban context. For a detailed discussion of development effects of polarisation, see Reynal-Querol et al. (2005).
Considering the Fractionalization index, Alesina et al. (2013) also observe that intermediate levels of diversity produced using this index can be the result of both a large pool of immigrants from only few countries of origin, or a small group of foreign-born hailing from a very diverse set of countries. To remedy this problem, they decompose the Fractionalization index into between- and within-components, where the former reflects the overall presence of foreign-born in an economy, and the latter captures the breadth of immigrant source countries. This ‘within’ component, which for convenience I call the Alesina index, is calculated as:

\[
Alesina_j = \sum_{r=2}^{R} \left[ \frac{n_j}{1-n_1} \times \left(1 - \frac{n_j}{1-n_1}\right) \right] \times (1-n_1)^2
\]

where \(n_1\) is the share of native-born workers in the city population (with other subscripts as above), and the equation is indexed over all nonnatives \((r = 2)\). This index is somewhat akin to estimating Fractionalization values among only the foreign-born population of each location.

It is instructive to compare how these indices differently characterize urban immigrant diversity in practice. Using data for U.S. metropolitan areas, built from an IPUMS five percent, five-year (2007-2011) extract of the American Community Survey (Ruggles et al., 2010), Table 1 reports the ten most diverse metros according to the Fractionalization index, and compares this with where these purportedly highly-diverse cities would rank using three alternative measures: the share of the foreign-born population, the Shannon index and the Alesina index. As expected, all of these measures are highly related, with the lowest correlation existing between the Alesina and Fractionalization indices (0.85, \(p = 0.000\)). The rankings in Table 1 reflect the strength of this relationship, but they also highlight potentially important differences. While the five or six most diverse cities according to the Fractionalization index appear in similar ranks across the other indices, the bottom half of the table presents a different picture. Made up of smaller places, chiefly in Texas, these are cities dominated by native-born Hispanics and Hispanic immigrants, as well as white, non-Hispanic natives. El Paso, for instance, located in West Texas and lying directly across the Rio Grande from Ciudad Juarez in Mexico, is 80% Hispanic and 15% white, with approximately one quarter of the population born abroad, overwhelmingly in Mexico. While, strictly speaking, El Paso is “composed of differing elements,” its situation does not closely match the notion of diversity that researchers or consumers of diversity research are likely to hold. It certainly contrasts with places like Miami, New York and San Francisco, in terms of the breadth
Table 1: Approaches to Measuring Immigrant Diversity

<table>
<thead>
<tr>
<th>CBSA</th>
<th>(I) Fractionalization Index</th>
<th>(II) % Foreign-Born</th>
<th>(III) Shannon Entropy Index</th>
<th>(IV) Alesina Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Miami-Fort Lauderdale-Pompano Beach, FL</td>
<td>0.601</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>San Jose-Sunnyvale-Santa Clara, CA</td>
<td>0.595</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Los Angeles-Long Beach-Santa Ana, CA</td>
<td>0.550</td>
<td>3</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>San Francisco-Oakland-Fremont, CA</td>
<td>0.501</td>
<td>6</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>New York-N. New Jersey-Long Island, NY-NJ-PA</td>
<td>0.484</td>
<td>9</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Salinas, CA</td>
<td>0.469</td>
<td>7</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>McAllen-Edinburg-Mission, TX</td>
<td>0.438</td>
<td>5</td>
<td>55</td>
<td>49</td>
</tr>
<tr>
<td>Laredo, TX</td>
<td>0.438</td>
<td>4</td>
<td>62</td>
<td>72</td>
</tr>
<tr>
<td>El Paso, TX</td>
<td>0.429</td>
<td>8</td>
<td>46</td>
<td>41</td>
</tr>
<tr>
<td>San Diego-Carlsbad-San Marcos, CA</td>
<td>0.410</td>
<td>12</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td>Correlations</td>
<td>1.000</td>
<td>0.995</td>
<td>0.953</td>
<td>0.854</td>
</tr>
</tbody>
</table>

(0.000) (0.000) (0.000) (0.000)

Note: n=285 metropolitan CBSAs. Correlation p-values in parentheses.
of nationalities present in significant quantities. Interestingly, El Paso and these other cities rank far lower on both the Shannon and Alesina indices, providing some support to the notion that these alternative measures capture diversity differently. Rankings using the Shannon and Alesina metrics place these seemingly less-diverse locations at more intuitively satisfying rungs in each index’s overall diversity ladder; the remainder of their most diverse cities include populous and highly mixed cities like Washington DC and Boston. Given the strong overall relationship between the Fractionalization and these other indices for U.S. cities, it is unclear whether these alternative measures will lead to different results in the kinds of empirical studies to be described below. But this kind of exploration is certainly warranted, given the Fractionalization index’s shortcomings in measuring diversity under the kinds of conditions actually found in most urban locations.

The next sections investigate the frameworks used and results produced in the empirical literature that investigates the links between immigrant diversity and productivity, innovation, and entrepreneurship.

4 Immigrant Diversity and Productivity

4.1 Analytical Framework

Approaching the relationship between productivity and immigrant diversity, research has mainly proceeded from economists’ spatial equilibrium tradition pioneered by Roback (1982), adapted by Ottaviano and Peri (2006; 2006) to the issue of immigrant diversity. In this model, a national economy is considered to be an open system composed entirely of a large number of distinct cities, in which firms and workers are free to choose locations. It is stipulated that workers’ utility – composed of wages, living costs and their consumption of location-specific amenities – will be arbitrated through mobility, thus tending toward equalisation over space. Immigrant diversity is conceptualized as a local public good, with potential feedbacks to both firm production capabilities and consumer satisfaction. Based on the theoretical discussion above, if immigrant diversity augments productivity, an exogenous increase in immigrant diversity in a particular city ought to increase wages, as workers are made more productive. It should also stimulate an increase in place-specific housing costs (commonly proxied using rents), as workers from other locations

4See Sparber (2008) for a different modelling approach.
crowd in seeking higher wages. To the extent that diversity curtails productivity, then wages and rents should decline. If diversity chiefly enhances consumer satisfaction, perhaps by offering a greater breadth of restaurants and other cultural amenities that residents value, then an increase in diversity could be associated with an increase in rents, as the location becomes more desirable, and a decrease in wages, owing to the fact that workers received benefits through these diversity-based amenities that reduce their need for higher wages. Meanwhile, if diversity represents a consumption disamenity, then an increase in diversity should reduce rents but increase wages, the latter as a means of compensating workers to endure the unpleasantness of living with people from different backgrounds.

The larger point is that, on the basis of the spatial equilibrium framework, one cannot infer from wages alone how diversity and productivity might be related. Instead, wage data can be usefully complemented with information about urban rental costs. On the basis of this logic, spatial equilibrium-flavored work has proceeded by jointly estimating rent and wage equations according to the following schematic:

\[
\begin{align*}
    w_j &= div_j + X_j' + \varepsilon_j \\
    r_j &= div_j + Z_j' + \varepsilon_j
\end{align*}
\]

(4) (5)

where \( w \) measures wages for workers in city \( j \), and \( r \) represents their rental costs. The primary predictor of interest is \( div \), indicating the city-specific level of immigrant diversity; \( X' \) represents a vector of city and worker characteristics to be analyzed that ought to affect wages; \( Z' \) is a set of characteristics that ought to shape rents, and \( \varepsilon \) is a disturbance term that satisfies classical regression properties.

Given the spatial equilibrium hypothesis’ strong assumptions around utility equalization, preference homogeneity, and firms’ ability to substitute among factors of production, it is certainly worth asking equally strong questions about the extent to which its identification strategies hold when one relaxes the assumptions. Indeed, some geographers have generated evidence that questions the logic of these models, specifically in terms of the forces guiding mobility decisions, and resulting implications for utility equalization (Kemeny and Storper, 2012; Storper and Scott, 2009). In the context of immigrant diversity, a few issues call this identification strategy into question.
Assuming a workforce with heterogeneous preferences, if particular kinds of workers prefer diverse environments, this may render shifts in wages and rents insufficient to identify how diversity is related to productivity. Similarly, certain firms’ specific needs of diversity or other forms of spatial preferences could limit researchers’ ability to cleanly observe the means by which diversity and economic outcomes are related. Nonetheless, though imperfect, examining diversity’s relationship with wages and rents provides a useful point of departure for the empirical study of diversity’s urban economic effects.

4.2 Empirical Results

Table 2 presents a selective summary of econometric evidence on the impacts of urban immigrant diversity. Ottaviano and Peri’s (2006) contribution is the initial paper in this field, and so a sensible place to begin. Examining diversity in 160 U.S. metropolitan areas, Ottaviano and Peri leverage public-use Decennial Census samples for 1970 and 1990. In order to control for differences in labor market composition, the authors restrict their sample to white native male workers between the ages of 40 and 50, and use average wages and rents for this subset of each city’s workforce as dependent variables. While they control for differences in the average educational attainment between each city’s native population, the ability to account for other sources of heterogeneity in the chosen worker subgroup is limited. The authors find that birthplace diversity, estimated using a Fractionalization index, is robustly and positively correlated with both wages and rents in both periods, suggesting that diversity acts chiefly to shift firms’ productivity curve outward. In their base specification, Ottaviano and Peri (2006) find that a rise in the Fractionalization index of 0.1, corresponding to the rise in immigrant diversity experienced in Los Angeles between 1970 and 1990, is associated with a 13% increase in native wages (and an even larger increase in rents). Potential for bias due to reverse causality is a concern: just as diversity may augment productivity, some immigrants’ will quite sensibly be drawn to strong economies. Using a shift-share measure of ‘predicted’ city-specific immigrant diversity, an approach that has been widely used in the broader immigration literature (see, for instance Card, 2001; Saiz, 2007), the authors perform two-stage least squares instrumental variables analysis. The results suggest that the direction of causality runs primarily from diversity to wages, and not the other way around. This study then suggests
that immigrant-induced diversity may have pronounced effects beyond the shifting of labor supply – diversity appears also to augment the wages of native workers living in its midst.

A host of studies follow the contours of Ottaviano and Peri’s approach, looking at other economies and time periods. Examining around 700 regions in 12 members of the EU15 in 2001, Bellini et al. (2013) find that birthplace diversity is positively associated with both gross domestic product (GDP) per capita and prices in ethnic restaurants – the latter proxying for amenities that ought to be capitalized in rents. In a study of 250 U.S. metropolitan areas in 2000, Kemeny (2012) also finds results suggestive of a positive productivity effect, measuring diversity’s links to both wages and employment, with employment included as a gauge of housing costs on the basis that increasing employment on fixed space will put upward pressure on the price of land. In terms of magnitudes, this study finds that a one standard deviation increase in diversity in U.S metropolitan areas could augment native wages by as much as 8%. Suedekum et al. (2009) also measure the relationships between diversity, wages and employment, examining 326 West German NUTS3 regions over the period 1995–2006 in a panel framework. In their overall sample, the authors find that diversity is negatively related to both wages and employment. Additionally, they distinguish between high- and low-skill native workers, on the basis that complex problems of the type for which identity and functional forms of diversity may be an advantage are not equally distributed in the working population; rather, they are more prevalent among workers who have higher skill levels. They find confirmatory evidence of the importance of this distinction: among high skill native workers, diversity is positively related to both wages and overall regional employment levels, while among less-skilled workers, diversity is negatively related to both outcomes. The authors argue that this finding at the low-skill end of the German labor market conceals some positive productivity effects at that level; these effects emerge only after accounting for a set of cities in which there is only a singular and large group of foreign-born low skill workers.

In the UK context, Nathan (2011b) examines the implications of migration and diversity over the period 1994 to 2008. The study finds a positive relationship between birthplace diversity and natives’ wages, and particularly so for skilled workers, with some negative findings for intermediate and less-skilled workers. Results on employment are positive for the more highly-skilled labor force and negative for those less-skilled, although the authors believe that this mostly reflects region-
<table>
<thead>
<tr>
<th>Outcome</th>
<th>Study</th>
<th>Unit of Obs.</th>
<th>Period</th>
<th>Diversity</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bellini et al. (2013)</td>
<td>EU NUTS3</td>
<td>1991, 2001</td>
<td>Birthplace</td>
<td>(+) Wages; (+) Rents</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Municipalities</td>
<td></td>
<td>Fractionalization</td>
<td>(-) Rents, IV robust</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Fractionalization</td>
<td>(insignificant in panel)</td>
</tr>
<tr>
<td>Output/Capita</td>
<td>Ager and Brückner (2011)</td>
<td>US County</td>
<td>1870–1920</td>
<td>Birthplace</td>
<td>(+) Output</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Alesina</td>
<td>IV robust</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Fractionalization</td>
<td>IV robust</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Herf. &amp; Shannon</td>
<td>IV robust</td>
</tr>
</tbody>
</table>
specific de-industrialization dynamics. In contrast to results for the U.S., Germany and selected EU15 regions, the study finds little significant relationship between diversity and the housing market among UK cities. In a cross-sectional analysis of British cities, Nathan (2011a) examines how a different definition of diversity based on the classification of names on the basis of cultural, ethnic and linguistic characteristics, is positive, but not significantly related to both wages and employment. Examining the Netherlands, Bakens et al., (2013) replicate the general approach taken in Ottaviano and Peri (2006), and find that a 0.1 increase in the birthplace Herfindahl is associated with a 9% increase in wages in the largest Dutch cities, and is robustly positively related to rents; however, the effect is negligible among all cities in the Netherlands. Sparber (2010) finds that racial diversity, with race defined as Asians; Blacks; Hispanics; Whites; and Other, is positively and significantly associated with wages and employment in U.S. metropolitan areas between 1980 and 2000, though the link to employment is insignificant. In the same analysis, he shows that racial diversity is unrelated to state-level changes in productivity.

Other work examines links between diversity and productivity measures, without considering the spatial equilibrium implications, in terms of housing prices, employment density or amenities. For instance, Ager and Brückner (2011) examine the historical peak of immigration to the U.S. – the period between 1870 and the start of the Second World War – during which the share of foreign-born in the U.S. population reached 15%. Using public-use data from the Decennial Census, the authors find that ‘cultural’ fractionalization – defined using a Fractionalization index calculated for a mix of national, regional and racial groups – is positively related to county output per capita. They also find that cultural polarization, meaning the situation in which a region contains both a large native majority and a singular large immigrant minority group, is negatively related to output.

In an international comparative study of 195 countries, Alesina et al. (2013) find that birthplace diversity is positively related to per capita GDP and total factor productivity. The strongest association exists among high skill workers, and in rich countries. Much of the international development literature has focused on ethnic fractionalization, to which the authors demonstrate that birthplace diversity is largely uncorrelated; while the consensus is that ethnic fractionalization inhibits development, these results suggest that immigrant diversity influences development through
a distinctive channel, generating much more positive economic outcomes.

Overall, the existing evidence investigating immigrant diversity’s relationship to productivity presents a fairly consistent picture. Most studies find that diversity is positively and significantly related to wages and either rents, employment, or amenities. These results are interpreted to indicate that the productivity-augmenting benefits of immigrant diversity outweigh the costs of transacting across cultures. Using instrumental variables techniques, much of the extant research has investigated the possibility of reverse causation, with scholars largely concluding that, in the observed relationships, the direction of causality runs from diversity to productivity and not the other way around. A number of studies also distinguish between the mostly positive effects of diversity as opposed to the negative impacts of polarization. Still, this work remains at an early stage; we are far from a point at which to confidently speak of a positive immigrant diversity productivity effect. The distance between the current state of our knowledge and this ultimate goal will be taken up in detail in section 7. The paper turns next to studies examining innovation, and then entrepreneurship.

5 Immigrant Diversity and Innovation

5.1 Analytical Framework

The research on immigrant diversity and innovation has taken a knowledge production function approach. Pioneered by Griliches (1979), the knowledge production function represents the idea that it is possible to explain levels of innovative output – whether at micro- or macro-levels – in terms of a set of innovative inputs, typically some gauge of research and development effort and human capital. As Audretsch and Feldman (2004) describe, in empirical tests of this idea, the relationships between such innovative inputs and outputs are considerably stronger at levels of aggregation above the firm, such as the industry or the country. Thinking about innovation in this sense, then, prompts the realization that it appears to have an extra-firm character, and considerable research indicates that the metropolitan scale plays a particularly important role in this regard (Jaffe et al., 1993; Rosenthal and Strange, 2008).

Based on the theorized mechanisms described in section 2, Niebuhr (2010) posits the following
where $I$ is a measure of innovative output for region $j$; $\tau(div)$ is a transaction cost that rises with diversity, such that $[1 - \tau(div)]$ represents the fraction of innovative inputs that can be used to produce $I$; $A$ describes existing technological knowledge; and $L$ represents the labor force, for which individual units vary only in terms of their nationality, $r$. Given this functional form, it is assumed that workers born in different countries represent economic complements in the production of innovations; immigrant diversity in the workforce will be positively related to innovative output with an effect size that is a function of the elasticity of substitution between national origins $1 - \alpha$.

While this particular formulation, which Niebuhr (2010) applies to R&D workers in German regions, is not canonical, it represents the approach most closely rooted in theory. Most work has employed the knowledge production function framework, but there is currently no widely accepted form. In contrast to the productivity-oriented work described in the previous section, analyses linking immigrant diversity and innovation have not sought to model implications for consumption, worker mobility or other spatial equilibrium conditions. In this respect, studies of this relationship fit more clearly within the economics of innovation, as opposed to equilibrium-flavored urban economics. The former mainly does not distinguish meaningfully between regional and national economies in terms of factor mobility, while for the latter, it is workers’ and firms’ mobility across the urban system that emerges as the primary equilibrating force.

### 5.2 Empirical Results

Although a great deal of empirical work explores the links between immigration and innovation, only a small fraction goes beyond aggregate measures of the share of foreign-born to consider actual immigrant heterogeneity. In the broader immigrant–innovation literature, Stephan and Levin (2001) observe that immigrants to the U.S. have disproportionately contributed to research and innovation in science and technology fields, and are highly over-represented among U.S. Nobel laureates. Examining a panel of 20 European national economies between 1995 and 2008, Bosetti et al. (2012) focus on the relationship between the share of foreign-born workers in highly-skilled occupations and patents applications as well as the total national scientific article citations. After
controlling for the existing national stock of scientific knowledge, they find that the national share of immigrants in these occupations is positively related to innovation outcomes. Similarly Hunt and Gauthier-Loiselle (2010) find that differences in U.S. states’ shares of foreign-born workers have a great deal of power in explaining differences in patents per capita, yet they find that this result can be explained not in terms of cultural explanations, but instead with recourse to immigrants’ greater educational attainment in science and technology fields. Shifting to the New Zealand context, Maré et al. (2013) consider how regional indicators of migrant presence condition a wide variety of innovation outcomes, from product introductions that are new to the world to entering an export market. After controlling for firm-specific capabilities, however, they find no remaining explanatory power for immigration. The literature on immigration and innovation (and entrepreneurship) has recently been reviewed by Kerr (2013).

At the firm level, some researchers have more explicitly explored the role of immigrant heterogeneity. For instance, using Dutch matched employer-employee data, Ozgen et al. (2011b) consider the relationship between birthplace diversity and innovation. They find that firms dominated by immigrants tend to innovate at a lower-than-average rate, while firms whose foreign workforces are highly diverse are more innovative, especially in terms of their propensity to introduce new products. Focusing on a sample of firms in the London metropolis between 2005 and 2007, Nathan and Lee (2013) find that those with ethnic- and birthplace-diverse owners and partners are significantly more likely to introduce new products and processes, though this form of diversity appears unrelated to commercialization activities. Moreover, they find that the positive relationship between diversity and innovation is not confined to firms operating in ‘high-tech’ sectors. Nathan (2013) assign inventors to ethnic groups on the basis of a name classification system, and finds that inventors patent more when they live in regional economies that feature a more ethnically-diverse pool of inventors. Parrotta et al. (2012) find that greater ethnic fractionalization in Danish firms is associated with both a greater number of applications across a greater breadth of patent categories; in this case diversity enters the regressions significantly only among white-collar workers. While these results are interesting, such studies consider causes and effects confined to the atomized firm, despite the abundant research indicating the importance of considering the geographical character of knowledge production.
Ozgen et al. (2011a) engage more directly with innovation’s regional dynamics. They find a modest positive relationship between regional birthplace fractionalization and patent applications per capita in 170 EU regions. In a study conducted on German planning regions, or Raumordnungsregionen, Niebuhr (2010) estimates birthplace fractionalization, as well as Shannon entropy indices, focusing on workers engaged in research and development (R&D) activities. Among inclusive samples and those restricted to only highly skilled R&D workers, and after controlling for major known determinants of innovation, this study concludes that regional immigrant diversity is positively related to per capita patenting activity.

Continued work on this topic is needed, especially conceptualizing diversity beyond simply capturing the extent of immigration.

6 Immigrant Diversity and Entrepreneurship

6.1 Analytical Framework

The link between diversity and entrepreneurship is rooted in the same theory of regionally-bounded knowledge spillovers as the other outcomes in this review, however, the mechanism is conceptualized slightly differently. Entrepreneurship has long been considered to be a function of individuals’ recognition of economic opportunities. Recent scholarship has sought to understand how differences in the presence of those opportunities may be systematically related to other economic decisions, in line with endogenous growth theory (Acs et al., 2009; Braunerhjelm et al., 2010). Under this view, entrepreneurship is a partly function of the availability of knowledge, which will be related to investments in knowledge production among incumbent firms and other knowledge-producing institutions. While knowledge production plays a role, agents must also recognize opportunity. In the process of recognition and valuation, diversity may be important: individuals from different backgrounds will frame and evaluate available knowledge differently, leading to a wider range of approaches to valuing and exploiting available ideas. This greater breadth of approaches ought to improve the likelihood that an idea will be exploited (Audretsch et al., 2010). At this juncture, there have been only a few attempts to model these dynamics. Researchers have used simple linear regression models, where rates of new firm births are used as the outcome of interest.
6.2 Empirical Results

Empirical work on the impact of diversity on entrepreneurship and employment is the least well explored of the topics discussed in this review. Three known studies exist at this time. Considering immigrant diversity at the firm-level in a panel of Danish firms, Marino et al. (2012) find that linguistic diversity is positively related to the likelihood that an individual will transition to self-employment. Taking a more explicitly regional approach, Audretsch et al. (2010) find that immigrant diversity, defined in terms of the Shannon index, is positively related to the share of startups in German planning regions. Diversity becomes highly significant in models estimated on technology-intensive sectors only. The results are robust in models that account for time-invariant unobserved regional heterogeneity, as well as those addressing bias from spatial autocorrelation. Focusing on firms in ten aggregate industrial sectors in U.S. counties, Cheng and Li (2011) take a Bayesian approach to examine the relationship between the formation of new, single-establishment firms and immigrant diversity, defined using the Shannon index. The results suggest that immigrant diversity plays little role in entrepreneurship in most of these sectors, except for ‘wholesale and retail’ and ‘leisure and hospitality.’ Diversity appears not to matter in sectors typically held to be sophisticated and information-rich, such as ‘professional and business services’ and ‘information,’ a finding that fits with results from a nationally-representative survey of immigrants and natives in fast-growing high-technology firms (Hart and Acs, 2011). Much more work on the links between diversity and entrepreneurship are needed to better understand the dynamics in this relationship.

7 Open Questions

Thus far, this review has focused on the available theory and evidence linking immigrant diversity and three economic outcomes: productivity; innovation and entrepreneurship. The remainder of the review is more prospective, concerned with highlighting the major open questions in the literature, and the kinds of approaches that might best answer them.
7.1 Endogeneity I: Heterogeneity and Sorting

The most sophisticated frameworks reviewed above endogenize workers’ and firms’ locational choices among cities in a larger national urban system. Yet, they do so by upholding some very restrictive assumptions about both workers and firms. For the purposes of tractability, the canonical Ottaviano and Peri (2006) model assumes that, but for their national origins, workers are identical. There are good reasons to develop models that incorporate a wider view of worker heterogeneity. For example, some researchers believe that certain locational amenities disproportionately attract highly skilled workers (Florida, 2002). Demand-side forces also have a role to play, as workers self-select into particular locations as they match their aptitudes and interests to the kinds of jobs available, which are a function of industrial mix (Combes et al., 2008; Kemeny and Storper, 2012; Moretti, 2013). If these sorting processes are orthogonal to inter-urban variation in diversity and productivity, then worker heterogeneity, as well as the sorting to which it gives rise, may not bias estimates of the impacts of immigrant diversity. However, this is unlikely to be the case. Consider Silicon Valley, which attracts workers with skills in computer science and related activities from other locations in the U.S., as well as from abroad. In this case, it is evident that sorting, diversity and productivity are deeply intertwined.

In theory, one might solve this kind of problem with sufficiently subtle data that could distinguish firms and workers on the basis of the kinds of characteristics that drive such sorting behavior. In practice, however, these data do not exist, nor are they likely to in the future. To get a sense of the problem, consider two white native male college dropouts, born in 1984: Mark Zuckerberg, residing and working in the San Jose-Sunnyvale-Santa Clara metropolitan area (the nexus of Silicon Valley), and Not-Mark Zuckerberg, who works in Omaha, Nebraska. Mark Zuckerberg’s locational choice of San Jose is clearly the result of a matching process between (a) the ensemble of industrial activities and social networks that are concentrated in the environs of Silicon Valley, and (b) his skill set in these fields, reflected in his productivity, innovativeness and entrepreneurial spirit; similarly Not-Mark has made locational choices that reflect a host of individual characteristics, including ones related to his productivity and that also likely pertain to the local industrial milieu. And yet, the canonical model proposes equivalence between workers who share age, race, nativity, and schooling characteristics. And in all likelihood, no amount of additional stratifying variables
will fully capture these kinds of important differences. We cannot differentiate between Mark and Not-Mark in the Ottaviano and Peri (2006) schema, with the possible result that the extant research has ascribed to immigrant diversity an effect that properly resides in hard-to-observe but nonetheless highly pertinent individual differences. In other words, Silicon Valley may be more immigrant-diverse than Omaha, while also paying higher wages, but, after accounting for differences in educational attainment and other factors, the residual wage gap may have much to do with Silicon Valley’s concentration of highly-skilled, innovative and driven workers, and little to do with diversity.

Though the Mark Zuckerberg thought experiment considers an extreme example, unobserved heterogeneity and associated sorting dynamics among natives and immigrants alike may play an important role in determining regional productivity and wages. Anecdotally, we know that regions like Silicon Valley lure the world’s best and brightest. On paper, many of these workers may look similar to other college-educated workers. It is likely that some of their high wages and productivity are due to in-situ learning, but it much of their economic value also lies in skills and aptitudes, and these characteristics are part of the reason they have settled in the region. Recent evidence from France, Sweden, the U.K. and the U.S. shows that fixed unobservable worker characteristics play an important role in explaining wage variation across urban systems, and that workers sort themselves, matching their skills to suitable localities (Yankow, 2006; Combes et al., 2008; Gibbons et al., 2010; Andersson et al., 2013). This sorting process appears to be especially important in jobs that intensively require cognitive and interpersonal skills (Bacolod et al., 2009; Andersson et al., 2013). Industrial structure could be the draw, as we can assume it is for Mark Zuckerberg. Heterogeneity in individual preferences for consumer amenities represents another potential factor driving sorting behavior. If highly productive workers disproportionately value Korean eateries, they may flock to Los Angeles, a city that boasts the largest Korean population outside Seoul. To the extent that these preferences cut across demographic differences, this will raise average levels of productivity among all subsets of native workers, but the effect has nothing to do with diversity’s purported performance-augmenting effects.

Two recent studies on urban immigrant diversity grapple with this potentially important issue. Examining the Netherlands, Bakens et al. (2013) exploit individual-level panel data on wages and
rents as a means to relax the assumption of individual homogeneity. To do so, they adopt a two-step approach proposed by Combes et al. (2008). In the first stage, the authors regress an individual’s wages on observable time-varying worker characteristics, a worker fixed effect, a sector-specific fixed effect, and a city-year fixed effect. Similarly, they regress each homeowner’s housing price per square meter on time-varying homeowner characteristics, dynamic home characteristics, a homeowner fixed effect, and again a city-year fixed effect. Results of this first stage provide an account of the relative importance of individual- versus city-specific drivers of wages and rental costs. To get a sense of the impact of diversity within the complex of city-specific factors, in the second stage, the authors use the coefficients on the first stage city-year fixed effect as outcomes in wage and rent equations. In other words, the second stage regresses the overall importance of city-specific factors in a given year on levels of urban immigrant diversity. Performing these stages, the authors find the following results: individual characteristics, both observed and unobserved, emerge as the primary determinants of variation in wages and rents. After considering these factors along with the sectoral effect, the sum of city-specific characteristics only modestly affect urban wages and rents. From their second stage regressions, they find that diversity is positively but largely insignificantly related to wages, while it is negatively and significantly related to rents. There is, in other words, little evidence that immigrant diversity augments worker productivity in Dutch cities, and this paper suggests that immigrant presence in cities reduces housing prices, rather than enhancing neighborhood quality. Whereas their estimates adopting the baseline model of Ottaviano and Peri (2006) produces results that are broadly in line with the existing literature, their conclusions are entirely different after considering worker heterogeneity; this ambitious paper raises provocative questions about much of the existing research.

Another way to address individual unobserved heterogeneity would be to exploit individual-level panel data on wages (and ideally rents) in order to estimate worker-level regressions, such that changes in a city’s level of urban immigrant diversity are related to changes in individuals’ wages and rents, with individual, city, time, industry and other fixed effects accounting for hard-to-observe stationary and dynamic factors. Just as the approach pursued by Bakens et al. (2013) is highly demanding in terms of data, so too is this method, for a few reasons. One is that estimating a model with so many fixed effects requires repeated measures on a very large number of individuals.
Another issue is that diversity levels may not be sufficiently dynamic over a given study period. Across many study periods, one might expect cities, and especially large cities, to have diversity levels whose rates of change are relatively sluggish. At minimum, researchers require large-\(N\), large-\(T\) panels to pursue this approach. A somewhat less demanding approach, akin to that taken by Gibbons et al. (2010) for Britain, would be to exploit individuals workers that move from one city to another across periods. This still requires longitudinal worker data, but it does not require a very long panel, nor does it rely on city-level dynamism in diversity levels. Given the right data, both of these approaches have the added benefit of capturing a dynamic relationship. Although Bakens et al. (2013) estimate a panel model in their first stage, they never address dynamics in diversity levels. Yet if researchers believe that diversity causes improvements in productivity, then we ought to be exploring models in which dynamics are built into the analytical approach. Such dynamic data offers clear econometric advantages, but it also permits better exploration of the hypothesized causal relationship. A working paper by Longhi (2011) is the only known paper to takes this kind of dynamic approach, using data on English workers in Local Authority Districts. She shows that, although cross-sectional results suggest a positive relationship between diversity and wages, panel estimate find no supportive evidence.

A second approach, also considering dynamics more explicitly, would exploit exogenous variation supplied by a natural experiment. This approach has been applied productively in related topics, such as the effects of changes in the supply of skilled immigrants (Kerr and Lincoln, 2010), and the links between ethnic enclaves and immigrants’ success in the labor market (Edin et al., 2003). In the context of the U.S., one possible source of such variation might be the 1965 Immigration and Nationality Act, or Hart-Celler Act. This law abolished a prior system of national quotas, resulting in a large increase in the breadth of immigrant source countries. Whatever the precise approach, however, researchers should pursue research designs that have a better chance of offering up an answer to the real question at hand: do changes in urban immigrant diversity result in improvements in productivity, innovation and entrepreneurship?

Worker heterogeneity and its implications for sorting is a major issue with which future research must wrestle. It is unclear what approaches like Bakens et al. (2013) would find for larger urban systems like that found in the U.S., but it seems plausible that they might conclude that current
estimates of the strength of diversity’s productivity impacts are overstated. Heterogeneity in firm locational choices may be equally important; it is even less frequently incorporated into spatial equilibrium frameworks. Progress on these issues could have major implications not just for work on the economic benefits of urban diversity, but for the broader fields of urban economics and economic geography.

7.2 Endogeneity II: Reverse Causation

The approach used to estimate immigrant diversity’s economic impacts leaves open the possibility that coefficients may be biased due reverse causation. While diversity might spur higher wages or greater levels of innovation, it seems likely that high native wages (and to a lesser extent innovation) will also attract immigrants, thereby raising the diversity index. Local diversity might thus be a function of local wages, as much as wage levels might be stimulated by the presence of immigrants with complementary capabilities. To the extent that this is true, regression estimates will overstate the exogenous impact of diversity on economic outcomes. As discussed earlier, recognising this issue, researchers have used instrumental variables techniques, with most papers using a variant of the shift-share instrument described in Card (2001), in which lagged city-specific shares of immigrant groups are combined with national growth rates of these groups to generate ‘predicted’ diversity levels for study years. Yet the validity of these instruments rests on the supposition that initial waves of immigrants chose locations based on extra-economic concerns. As Aydemir and Borjas (2011, p.30) observe, “If the earlier immigrant arrivals selected those markets because they offered relatively better job opportunities, any serial correlation in these opportunities violates the orthogonality conditions required in a valid instrument.” Hence, we are far from a place at which secure statements can be made about the direction of causality in the relationship between diversity and economic outcomes, and therefore also in the size of effects running in any particular direction. To improve this situation, new instruments needs to be devised, alongside quasi-experimental approaches described in the previous section.
7.3 Labor Market Segmentation and Differential Effects

Immigrant diversity’s economic impacts may vary among different segments of the labor market, and this topic has been taken up in recent research, with efforts focused chiefly on productivity outcomes. Diversity could enhance the productivity of certain workers while making other workers worse off. Foreign-born workers may displace natives at the lower-skill end of the labor market, though they may also complement them (Borjas, 2003; Peri and Sparber, 2009; Peri, 2012). One way to interpret the theoretical discussion in Section 2 is that, because highly-skilled workers are those whose jobs involve the greatest complexity and problem-solving, it is these workers for whom productivity effects due to diversity-induced complementarity will be strongest (Hong and Page, 2001; Weber and Fujita, 2004). Other axes of differentiation are also possible. For instance, Sparber (2010) posits that diversity enhances ‘creative’ industries, where interaction may fuel ideas but is not needed for implementation, while reducing productivity in sectors that depend more on ongoing group participation. Industries represent another possible way that immigrant diversity may have a differentiated effect. Skills, tasks and industries may be highly correlated in some situations and largely unrelated in others.

Existing studies exploring potential differential effects of urban immigrant diversity have so far distinguished between worker groups on the basis of educational attainment. Initial results confirm that observed positive productivity effects of immigrant diversity for highly educated workers are stronger than for those for workers with relatively less schooling (Suedekum et al., 2009; Nathan, 2011a). These papers consider how diversity affects workers with different levels of educational attainment, but immigrants themselves are also differentiated on the basis of their educational attainment. Investigating this idea, Alesina et al. (2013) build separate national measures of diversity for more- and less-educated immigrants, and find an especially strong association between average wages and a diverse mix of highly-skilled immigrants. Applying this approach to the regional scale using data for U.S. metropolitan areas, Alesina indices (equation 3) are constructed separately for (a) immigrants with at least a Bachelor’s degree, and (b) those who have not completed secondary school. Figures 2 and 3 display the bivariate relationship between these high- and low-skill immigrant diversity measures and average metropolitan wages among white native male workers between the ages of 40 and 50. The first point to make about these indices is that they
diverge from those that do not heed differences in immigrant educational attainment. This is particularly true for the index capturing diversity within the category of highly-educated immigrants, whose average correlation coefficient with the skill-undifferentiated diversity measures reviewed in Table 1 is 0.6. The low-skill measure remains closely related to with the prior measures. Among other things, these differences suggest that skilled immigrants are less evenly distributed among cities in the overall immigrant population. Differences in the distribution of skilled immigrants is reflected in the rankings of ten cities that are most diverse according to each measure, which are labelled in Figures 2 and 3. Immigrants with high levels of educational attainment sort themselves into particular locations in the U.S. urban system, settling especially in its largest cities, as well as smaller urban areas like Trenton-Ewing and Ann Arbor that feature prominent research universities (Princeton and the University of Michigan, respectively). Communities that feature the greatest diversity among less-well-educated immigrants include some of same major metropolises, alongside smaller cities in California that are marked by very large relative quantities of Mexican and other Hispanic immigrants. The second important point is that, though both are positively related to natives' wages, the strength of their bilateral relationship differs considerably. Specifically, the slope of the line is much steeper for highly-skilled immigrants. Although these simple correlations should not be read as evidence of a causal relationship, they do point to the potential for differentiated effects that fit with existing theory. Rigorous analytical frameworks are needed to explore whether these differences truly matter for economic outcomes. Future studies should also consider other axes by which the effects of diversity may vary, from task dynamics to industries.

One challenge to identification among less-skilled workers is the possibility that any productivity or innovation effects may not be observed in wages nor in innovation indicators like patents. As Autor et al. (2008) describe for the U.S. (with similar trends for Britain documented by Goos and Manning (2007)), wages for workers at middle and lower segments of the labor market have grown much less rapidly than for those at the top end of the distribution. While some degree of this differential growth is certainly rooted in differences in productivity growth, some portion of it is likely due to the erosion of unions and other labor market institutions, as well as other factors. It is possible, therefore, that some degree of production gains in middle- and lower-tiers of the labor market may not be reflected in these workers’ wages. Meanwhile, the kinds of in-
Figure 1: Native Wages and High-Skill Immigrant Diversity

Figure 2: Native Wages and Low-Skill Immigrant Diversity
novations produced in these segments may not result in patents. Examining the very bottom of the labor market, Iskander et al. (2010) provide a case study of undocumented Mexican construction workers in Philadelphia, finding that these workers blended Mexican and U.S. construction techniques to produce process innovations that enhanced renovations of existing row homes. Yet these innovations did not raise their wages, due to a combination of their precarious labor market position and the end of the aughts’ construction boom; moreover, this new knowledge could not be manifested in data on entrepreneurship, since laws prevent undocumented workers from starting new businesses. Though an extreme example, it illustrates the need to look beyond wages and traditional measures of innovation and entrepreneurship in order to explore possible economics gains due to diversity at lower segments of urban labor markets. Case studies like Iskander et al. (2010) are a fruitful approach, yet ideally one would like to complement such methods with larger-scale econometrics, in order to get a sense of the significance and magnitude of potential economic effects. But such statistical approaches must confront the considerable challenges of capturing the outcomes of interest.

7.4 Institutions and Other Moderators

Given theory’s description of a double-edged relationship between diversity and economic performance, there is a need to understand the circumstances that determine when immigrant diversity is a benefit and when it is a hindrance. Scholars examining the international context posit that institutions, by reducing inter-personal and inter-group transaction costs, can minimize rent-seeking behavior and therefore improve economic performance under conditions of diversity (Costa and Kahn, 2003; Collier and Gunning, 1999; Alesina and La Ferrara, 2005; Easterly, 2001). Much of this research has focused on diversity, and especially linguistic and ethnic diversity, in the context of developing economies. Collier (2000), for instance, finds that ethnic diversity is unrelated to growth in developing countries that have democratically-elected governments, while it is negatively

5Indeed, a wealth of evidence suggests that such transaction costs are likely to be particularly high in culturally-diverse environments (Putnam, 2007; Leigh, 2006; Alesina and La Ferrara, 2002, 2000; Costa and Kahn, 2003; Collier and Gunning, 1999). These findings appears to be robust, at least in the short run, at scales ranging from U.S. counties to national economies, although the relationship is clearest when considering racial, as distinct from cultural fractionalization. Although this suggests that trust and diversity are endogenously related, this is not the same as saying that trust is mechanically determined by diversity: some degree of trust is likely determined by forces that are exogenous to diversity – a byproduct of other institutions and historical circumstances.
associated with growth in those economies that are run by dictatorial regimes. But, it is not implausible that institutions could moderate the effects of diversity in other contexts, including in metropolitan regions within advanced economies. Though American cities will not vary in terms of the architecture of their national political economy, formal and informal institutional structures can vary considerably within national boundaries, with significant implications for industrial and regional performance (Saxenian, 1996; Storper, 1997). The relevant regional institutions may be those that foster residents’ belief that their own well being, and that of their specific cultural group, is related to the welfare of other groups in their community. Formal mechanisms that induce inter-group interaction, such as organizations that seek to aid immigrants in integrating in their communities, are one example of such institutions. Spending on public goods such as housing and education may also signal diversity-friendly institutions, as may higher minimum wage regulations. Relevant informal institutions would include attitudes toward group boundary-spanning and reciprocity implicit in Granovetter’s (1973) notion of ‘weak ties,’ or what Putnam (2000) calls ‘bridging’ social capital. Together, these institutions could help agents harness the productivity-enhancing externality that is latent in diversity.

Kemeny (2012) explores this notion, finding that the positive relationship between diversity and wages found by Ottaviano and Peri (2006) and others depends on the degree to which regions have sufficient social capital. Using data on metropolitan areas for 2000, this study finds that native workers in highly immigrant-diverse cities that feature high levels of social capital earn seven percent higher wages than their counterparts living in equally diverse locations that have low levels of social capital. For workers in cities with weak institutions, the effects of diversity are nearly zero. The only other known study to address topics of this sort is Alesina and La Ferrara (2005), who investigate whether the effect of diversity on population growth in U.S. counties depends on whether a county is initially rich or poor. They argue that economies need robust institutions to cope with diversity, and assume that initial economic welfare is a reasonable proxy for those forces. They find that counties that are both poor and ethnically heterogeneous experienced negative population growth, while in wealthy, diverse counties, this negative relationship is weakened, and

---

6One might reconcile development economics’ chiefly negative view of diversity with economic geography’s mostly positive one with reference to differences in specialisation patterns: in poor countries, economic activities are less oriented toward complex problem solving; instead they are routine, providing less scope for the kinds of intellectual cross-pollinations that drive alleged benefits in advanced regional economies.
in some specifications, reversed.

Given the wealth of theory and evidence supporting the role of institutions – broadly conceived to include both formal and informal social structures – in economic affairs, and precisely in affecting inter-group transaction costs, much more research is needed into the social factors that may determine the circumstances under which urban immigrant diversity can bear economic fruit. Formal institutions represent a sensible avenue for further study, with investment in public goods reflecting a commitment to interaction and tolerance.\(^7\) Spatial patterns of work and habitation are also potentially important moderating forces, being both an outcome of preferences toward intercultural interaction, as well as a determinant of potential for interaction. It must be said that the literature reviewed in this paper has studied metropolitan areas as though they had no internal spatial structure that could enhance or inhibit interaction. And yet, alongside the rise in urban diversity, in the U.S. at least, has come a resurgence in immigrant spatial segregation, even as racial segregation has declined (Cutler et al., 2008). Segregation in immigrant gateway cities is particularly pronounced (Clark and Blue, 2004). In many American urban centers, particularly those that flourished after widespread automobile ownership became the norm after World War Two, it is entirely possible for sufficiently affluent residents to “isolate themselves from people of other cultures via the buildings they live in, the schools to which they send their children and their use of private automobiles rather than public transport” (Storper and Manville, 2006, p.1256). Rather than assimilating in both spatial and social senses, immigrant groups, meanwhile, remain segregated in ethnic enclaves. In Europe, segregation levels are substantial, though lower than in the U.S., and differences in assimilation in spatial and social terms are related in part to the generosity of welfare programs, and to state orientations toward multiculturalism (whether melting pot or mosaic) (Musterd, 2005; Koopmans, 2010). However, the larger point is that for cities of all kinds it seems entirely sensible that intra-urban settlement patterns among native and immigrant groups should moderate any relationship between diversity and economic performance, such that greater segregation ought to reduce immigrant diversity’s positive productivity or innovation effects. This could be the result of direct inhibition of potential interactions, as well as indirectly through a sorting process by which individuals with little taste for inter-cultural interaction choose

\(^7\)The available evidence for the U.S. suggests that spending on such goods is inversely related to diversity (Alesina and La Ferrara, 2005)
highly segregated locations. Its relationship to diversity as a consumer good is less clear: residential segregation may do little to dampen the ability of natives to sample products and services borne of diversity, with cross-cultural interaction occurring only in service capacities. Problematically, segregation and internal spatial structure remain entirely unexplored in this literature; the field needs to better understand these and other moderating relationships, as they have real potential policy implications.

7.5 Immigrant Diversity: Interdisciplinary Dialogue, Scale and Mechanisms

As has been discussed in depth above, economic geographers mostly assume that diversity is a metropolitan public good that generates positive or negative production externalities. In contrast, psychologists, organizational researchers, and other scholars contend that the critical interactions between birthplace-diverse individuals occur within organizations and work teams (Williams and O’Reilly, 1998; Webber and Donahue, 2001). Both presumptions are reasonable and supported by theoretical priors; indeed external and internal economies of immigration diversity may co-exist. Yet, at present, economic geographers have barely dipped their toes in the vast sea of papers and books on workgroup diversity in cognate subfields. Indeed, much urban-focused scholarship appears to be only superficially aware of the organizational literature. For instance, despite seminal experimental workgroup papers like Hoffman and Maier (1961) and Triandis et al. (1965), and the wealth of theoretical and empirical papers that have followed them, Niebuhr (2010, p.564) still refers to this literature as “emerging.”

There are several very good reasons for seeking to foster better integration between geographers and organisational researchers. One is that the links between these two open up an interesting research question, and one with considerable practical importance: are the economic effects of immigrant diversity (if any) concentrated in interactions within workgroups and organizations, or are they better understood, as is the knowledge production function, at a more aggregate spatial scale? Trax et al. (2012) and Lee (2013) represents the only known attempts to answer this question. Estimating production functions for a panel of German plants, Trax et al. (2012) consider how both plant- and city-level immigrant diversity, measured using Fractionalization indices, may affect total factor productivity. Controlling for the share of foreign-born, they find
that both plant- and regional immigrant diversity are positively, significantly and importantly related to plant productivity, with effect sizes in manufacturing for each manifestation of diversity in the area of 10 percent for a one standard deviation increase in heterogeneity; in other words, the magnitude of the relationships are large. Lee (2013) contrasts the effect of diversity among managers in a panel of small- and medium-sized UK enterprises and the overall diversity of the cities in which they operate, finding a modest firm effect but no city effect. However, diversity is here defined as the share or workers or managers who are born abroad. Further study of this topic is needed, in order to adequately understand the context in which diversity can lead to innovation, entrepreneurship and improved economic performance, and the relative importance of city- and establishment- or firm-specific manifestations of immigrant diversity in relation to such outcomes.

A second reason, and a more fundamental one, is that the conclusions reached in the urban and workgroup literatures appear to be quite different. The field of workgroup studies is large enough to have been the focus of several meta-analyses (Bowers et al., 2000; Webber and Donahue, 2001; Stewart, 2006; Horwitz and Horwitz, 2007; Hülsheger et al., 2009; Joshi and Roh, 2009). The results of the decades of workgroup studies that these meta-analyses synthesize are not particularly encouraging for urban-focused investigations of this topic. Surveying their field, reviews generally find that, despite the wealth of studies on team diversity, its links to team performance remain unclear in terms of significance, magnitude and direction. Team diversity has at best little consistent and powerful relationship to group effectiveness, problem-solving, creativity and innovation; where positive effects are noted, they are modest.

The disjuncture between weak organization-level findings and positive and fairly strong regional-level results may have a few potential explanations. First, it could be a function of the different ways that researchers in each field operationalize the notion of diversity. Organizational work has commonly defined diversity in term of age and gender, with a much smaller interest in nationality. So it is conceivable that while heterogeneity in workgroup ages and gender may be unrelated to performance, team immigrant diversity could be a more powerful wellspring of productive heterogeneity. In the team literature, diversity is also typically measured differently, given that heterogeneity among small teams cannot be effectively described using standard fractionalization measures. It could also be that, at the team scale, diversity is associated with transaction
costs that largely cancel out any positive effects, but that these costs are muted at the urban scale (though it is not obvious why this might be the case). There is no clear way to bridge these contrasting results, in no small way because urban studies have yet to overcome some of the issues raised in the sections above.

Additionally, urban studies face the challenge that their measure of immigrant diversity reflects instead an omitted variable. This is, in a sense, built into operationalizations of these models: it is assumed that one’s birthplace indicates in some meaningful way one’s manner of approaching the world; that “bio-diversity” reflects intellectual diversity, the latter being latent and too hard to capture at sufficient scale. That diversity may proxy for a latent variable is itself a fatal problem, as long as there are not likely to be other latent variables that are also strongly correlated with diversity, and which affect productivity, innovation and entrepreneurship. This is a big assumption, and at least a few challenges to it are worth making.

Consider, for instance, diasporic networks, which represent a related channel through which diversity can positively influence productivity, whether in work groups or regions. Research suggests that immigrant networks could reduce transaction costs and promote the exchange of goods, factors of production and ideas (Rauch, 2001; Combes et al., 2005; Saxenian, 2007). If productivity growth is chiefly a function of an economy’s ability to innovate and adopt existing technology, connections that facilitate global idea-sharing could stimulate productivity growth (Benhabib and Spiegel, 2005; Docquier and Rapoport, 2012). In the context of the research topic at hand, this channel is really a subset of the larger mechanism relating diversity of productivity gains described above. In the primary case, it is assumed that the value of diversity, in terms of enlarging the range of available solutions in the problem space, is rooted in birthplace-specific heuristics and perspectives. Another possibility is that, rather than some inbuilt culture-specific characteristics, foreign-born individuals enjoy international social connections to which natives lack access. Reduced transaction costs in these networks could aid the spread of valuable ideas. Most of the literature on this topic has focused on the effects of these diasporic networks on developing economies that have witnessed a ‘brain drain’ to advanced economies like the U.S. To the extent that organizations in advanced economies sit at the global technological frontier while those in developing economies lag behind, this focus on developing economies makes sense: it is in these less
developed economies that the effects of pipelines of knowledge and the dissemination of global best practices will be strongest. Existing evidence confirms the existence of this effect. For instance, analyses of patent citations indicate that brain drains from India and China to the U.S. have resulted in knowledge flows back to the country of emigration (Kerr, 2008; Agrawal et al., 2011). Case study research confirms this result (Gaillard and Gaillard, 1997; Saxenian, 2007). There is very little systematic evidence regarding the relationship in reverse; we do not yet understand the extent to which diasporic networks in diverse organizations have enhanced performance in organizations or in advanced regional and national economies. Given that, using standard quantitative approaches, these effects are likely to be indistinguishable from positive effects due to identity diversity, case studies and other empirical research methods will likely to be required to disentangle the relative importance of these potential channels.

Consider also the economic literature describing demand-side effects of immigration. Borjas (1995), for instance, describes as model in which immigration can augment demand for native factors of production to the extent that immigrants complement but do not substitute for those factors. To the extent that this mechanism operates in the economy, we would like to distinguish its positive effects from those that arise due to urban (or firm-specific) intellectual cross-pollination. Yet, as shown in Section 3 above, the commonly used indices to capture urban immigrant diversity are nearly perfectly correlated with city-level measures of the share of foreign-born. Immigration-spawned complementarity is likely to be an urban phenomenon. Researchers therefore face a seemingly irreducible mechanical confounding problem.

On top of this, consider that country of birth may be a poor proxy for functional diversity, the latter defined in terms of the ways that problems are perceived and solved. By dint of national culture, is it plausible that all Brazilian-born emigrants living in a particular economy share a certain manner of approaching particular problems and issues? This fails the test of common sense. And to the extent that such national regularities might exist, would they be the ones that matter for economic performance? Much of the current organizational literature on diversity has now shifted away from identity constructs like gender, age and birthplace, toward measures that seek to get closer to the hypothesized driver of economic rewards: diversity in terms of perspectives and heuristics (Dawson, 2012). The appeal of demonstrating positive effects of immigration in cities
is clear. But as social scientists, the primary goal must be to improve our understanding of the underlying mechanism. The idea that national culture shapes heuristics and perspectives ought to be subject to empirical validation. And to the extent that the important sources of heterogeneity lie elsewhere, researchers may be better off exploring how urban variation in heuristics, however derived, affect economic outcomes of interest.

8 Conclusion

Immigrant-induced diversity is growing in many metropolitan areas around the world, especially cities in North America and Europe. This paper has considered the empirical, and mainly quantitative literature on the economic effects of this diversity. This research is motivated by a theorized mechanism, developed chiefly in the study of workgroups and organizations, by which diversity may augment performance by increasing the range of ideas to be employed in problem solving, and at the same time may generate frictions that render co-operation more costly. Though this mechanism has been the focus of a very large literature at the level of workgroups and organizations, it is only in the last decade that economic geographers have transposed these ideas to the space of the city, in order to investigate immigrant diversity’s links to metropolitan economic performance.

While academically interesting, this research also offers potentially large policy implications. In policy as well as popular debates over immigration and the culturally-complex environments it engenders, it is mostly assumed that foreign-born workers displace natives. Rising immigration is thus cast as a zero-sum game. There is much less public understanding that immigration affects not just labor supply but also labor force composition, particularly in metropolitan areas. By exploring the potential effects of this shift in composition on the production process (as well as consumption factors), this literature has the potential to shift the terms of debate by showing that immigration-induced diversity may increase the economic welfare of natives and foreign-born alike.

To date, this is precisely what much of the empirical work suggests, finding a strong positive association between diversity and productivity, innovation and entrepreneurship in the U.S. as well as a number of European countries. This positive relationship appears largely robust to potential bias from reverse causation that could arise as mobile immigrants select into already strong urban economies, and is reasonably consistent across a range of different specifications with different
control variables.

Yet many questions remain about the validity of the findings of this growing literature. By now we have mostly picked the low-hanging fruit, by assuming workers are homogenous but for their national origin. One of the largest challenges to this body of scholarship is the fact that workers are indeed different, both in terms of the kinds of urban amenities they like to consume, and in terms of their productivity, innovativeness and entrepreneurial capabilities; most problematically, many of these differences are hard to directly observe. Yet they are at the base of a sorting process among cities that has much to do with aggregate levels of wages, rents, and a wide range of other measures of economic performance. Future work must dedicate itself to developing approaches that account for this sorting behavior and its consequences. This work will require both creative approaches, as well as, most likely, excellent data that permit tracking individuals over time. Equally important, it is not yet clear that heterogeneity of birthplaces effectively captures important kinds of variation in functional diversity that theory predicts should be the resource from which economic benefits are drawn; while at the same time, measures of immigrant diversity are likely to be highly correlated with other potential drivers of economic advantages. Other important open issues include potential differentiated effects of diversity on the basis of skills, tasks and industries; the role of institutional and spatial moderators; and the relative importance of city- and organization-specific manifestations of diversity. By addressing these challenges and gaps in the work to come, we will better understand how (and if) our increasingly immigrant-diverse urban environments shape economic performance.
References


46


