New Frontiers in Regional Science: Asian Perspectives 40

Rachel S. Franklin *Editor*

Population, Place, and Spatial Interaction

Essays in Honor of David Plane

Springer

New Frontiers in Regional Science: Asian Perspectives

Volume 40

New Frontiers in Regional Science: Asian Perspectives

This series is a constellation of works by scholars in the field of regional science and in related disciplines specifically focusing on dynamism in Asia.

Asia is the most dynamic part of the world. Japan, Korea, Taiwan, and Singapore experienced rapid and miracle economic growth in the 1970s. Malaysia, Indonesia, and Thailand followed in the 1980s. China, India, and Vietnam are now rising countries in Asia and are even leading the world economy. Due to their rapid economic development and growth, Asian countries continue to face a variety of urgent issues including regional and institutional unbalanced growth, environmental problems, poverty amidst prosperity, an ageing society, the collapse of the bubble economy, and deflation, among others.

Asian countries are diversified as they have their own cultural, historical, and geographical as well as political conditions. Due to this fact, scholars specializing in regional science as an inter- and multi-discipline have taken leading roles in providing mitigating policy proposals based on robust interdisciplinary analysis of multifaceted regional issues and subjects in Asia. This series not only will present unique research results from Asia that are unfamiliar in other parts of the world because of language barriers, but also will publish advanced research results from those regions that have focused on regional and urban issues in Asia from different perspectives.

The series aims to expand the frontiers of regional science through diffusion of intrinsically developed and advanced modern regional science methodologies in Asia and other areas of the world. Readers will be inspired to realize that regional and urban issues in the world are so vast that their established methodologies still have space for development and refinement, and to understand the importance of the interdisciplinary and multidisciplinary approach that is inherent in regional science for analyzing and resolving urgent regional and urban issues in Asia.

Topics under consideration in this series include the theory of social cost and benefit analysis and criteria of public investments, socio-economic vulnerability against disasters, food security and policy, agro-food systems in China, industrial clustering in Asia, comprehensive management of water environment and resources in a river basin, the international trade bloc and food security, migration and labor market in Asia, land policy and local property tax, Information and Communication Technology planning, consumer "shop-around" movements, and regeneration of downtowns, among others.

Researchers who are interested in publishing their books in this Series should obtain a proposal form from Yoshiro Higano (Editor in Chief, higano@jsrsai.jp) and return the completed form to him.

More information about this series at http://www.springer.com/series/13039

Rachel S. Franklin
Editor

Population, Place, and Spatial Interaction

Essays in Honor of David Plane

Springer

Editor
Rachel S. Franklin
Center for Urban and Regional
Development Studies (CURDS)
Newcastle University
Newcastle upon Tyne, UK

ISSN 2199-5974 ISSN 2199-5982 (electronic)
New Frontiers in Regional Science: Asian Perspectives
ISBN 978-981-13-9230-6 ISBN 978-981-13-9231-3 (eBook)
https://doi.org/10.1007/978-981-13-9231-3

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Introduction

Among the many ways to celebrate an academic career—parties, cake, conferences, or symposia—the *Festschrift* stands out. Communication, after all, is a hallmark of scholarly activity, and the written word holds a particularly special place in the pantheon of scholarly interaction. It is through the written word that knowledge is not only shared but also preserved. The production of a *Festschrift* illuminates additional aspects of an academic career, however. The nature of *who* contributes and *what* they contribute says so much about the person being honored—the range of topics addressed, as well the network of students, colleagues, and friends that emerges from the authors listed in the table of contents. In the end, the *Festschrift* commemorates the entire career: the research, the teaching, the mentorship, and myriad other forms of academic legacy.

The person celebrated here is David Plane, geographer, regional scientist, and demographer—but also coauthor, colleague, advisor, mentor, and friend. Our discipline would not be the same without him. Certainly, his scholarly contributions are remarkable in quantity and quality, but also for the relatively rare virtue of distinctiveness. For one thing, his work is instantly identifiable, notable for his long-standing inquisitiveness around spatial aspects of migration and population distribution. His papers have also often stood out for a playfulness and humor not often encountered in academic writing. His 1993 *Geographical Analysis* paper, "Requiem for the fixed-transition-probability migrant," in conversation with research published by Andrei Rogers, stands out, as does his 1991 book chapter, "The Ten Commandments of Migration Research," written with Peter Rogerson. At the same time, his work perfectly integrates into multiple disciplines at once, whether demography, regional science, or geography. His research has been published in the flagship journals of all three disciplines, and, chameleon-like, he engages with the questions, methods, and data of each as if it were his primary academic field.

Looking back, his career has been characterized by an interesting juxtaposition of continuity and continual change. For example, with the exception of sabbaticals (and his stint as a statistician at the US Census Bureau before moving to Arizona), the entirety of Dave's professional life was spent at the University of Arizona in Tucson. A strong thread of continuity runs through his research as well: largely analytical inquiry about migration flows, migrant characteristics, and the relationships between origins, destinations, and the bigger picture of population distribution in the United States. Further consideration of his research, however, reveals an almost omnivorous and continually evolving approach to methods and models. The topics—migration or demographic change or population distribution—might stay the same, but the tools used to elicit information about the phenomena at hand are often new. This is not a researcher with one tool in his toolbox (and only one craft to ply) but rather a magpie of methods. His research displays, as well, an affection for descriptive analysis, conceptual underpinning, and narrative. The methods and data employed in his research are not always simple, but the explication of results is typically remarkably elegant and straightforward.

There is another element to Dave's research that is nicely highlighted by this *Festschrift*, and that is his ongoing collaborations with colleagues and former students. His is a loyal and stalwart character. His enduring research partnership with Peter Rogerson is well-known, spanning as it does one textbook, *The Geographical Analysis of Population* (1994), and numerous scholarly articles. He has also published widely and generously with colleagues, including Beth Mitchneck, Gordon Mulligan, Daoqin Tong, and Brigitte Waldorf, and students, among them myself and Jason Jurjevich (but also many others). When presented the opportunity, all were honored to contribute chapters to this volume. Moreover, all would be pleased to call him friend.

The contributions contained herein, and the authors who so graciously took part in this endeavor, provide a window into the sorts of research topics that not only attracted Dave's interest but have piqued that of so many other scholars in the field. There are methods for population distribution (Rogerson, Chap. 1) and capturing income effects of migration (Ledent, Chap. 2, and Newbold, Chap. 9). There are other methods addressed, too, for spatial spillovers (Chung and Hewings, Chap. 3), location modeling (Tong, Mu, and Li, Chap. 4), and measuring human capital (Franklin, Chap. 5). There's engagement with the recurring issue of survey data quality, a topic that has often captured Dave's attention (Jurjevich, Chap. 7). And then there are many applications of migration and residential mobility, often with an explicit nod to Dave's research on age articulation or the relationship between labor markets and migration (Mulligan, Nilsson, and Carruthers, Chap. 6; Clark and Lisowski, Chap. 8; Raymer, Liu, and Bai, Chap. 10; Alimi, Maré, and Poot, Chap. 11; and Kim and Waldorf, Chap. 12). There's also investigation of the population-environment connection, with Gober's (Chap. 13) consideration of water use and demographic change. Finally, Mitchneck (Chap. 14) helps conclude the volume with a rumination on research collaboration (in particular, her work with Dave) and impact.

One last point: along with his research and mentorship, Dave is equally known for his service, especially in regional science, as the long-time executive secretary of the Western Regional Science Association (WRSA) and his activity in the Regional Science Association International (RSAI)—appropriate, perhaps, for one with a PhD in Regional Science. Foremost among his service activities have been those as a journal editor, of both *Papers in Regional Science* and the *Journal of Regional*

Science. Those who have worked with him and know him well will be aware how seriously he takes the role of editor—not only content but style, punctuation, and grammar. Thus, you can imagine the trepidation of this editor, trained by Dave no less, in making a gift to him of this volume. Forgive the occasional misapplied em or en dash, and please accept this offering in the spirit in which it is intended: as a commemoration of an illustrious academic career and as a token of friendship from all those represented herein.

Centre for Urban and Regional
Development Studies (CURDS)
Newcastle University
Newcastle upon Tyne, UK

Rachel S. Franklin

References

Plane DA (1993) Requiem for the fixed-transition-probability migrant. Geogr Anal 25: 211–223

Plane DA, Rogerson PA (1991) The ten commandments of migration research. In: Boyce DE, Nijkamp P, Shefer D (eds) Regional science: retrospect and prospect. Springer, Berlin, pp 15–42

Plane DA, Rogerson PA (1994) The geographical analysis of population: with applications to planning and business. Wiley, New York, 417 pp

Contents

Contributors

Omoniyi B. Alimi is teaching fellow in the School of Accounting, Finance and Economics at the University of Waikato in Hamilton, New Zealand.

Xujing Bai is a research assistant in the School of Demography at the Australian National University.

John I. Carruthers is an associate professor of Regional Science in the Department of City and Regional Planning at Cornell University; his Ph.D. is from the University of Washington.

Sungyup Chung is a research economist at the Bank of Korea in Seoul currently on secondment at the European Central Bank in Frankfurt, Germany. He completed his Ph.D. from the University of Illinois at Urbana-Champaign.

William A. V. Clark is research professor of Geography at the UCLA. He has had a career-long interest in demographic change in large urban areas and has published extensively on demographic change, models of residential mobility, and sorting processes that bring about residential segregation in the urban mosaic.

Rachel S. Franklin is professor of Geographical Analysis in the Centre for Urban and Regional Development Studies (CURDS) and the Spatial Analytics and Modelling Lab at Newcastle University in the United Kingdom.

Patricia Gober is research scientist and professor emeritus in the School of Geographical Sciences and Urban Planning at Arizona State University and in the School of Public Policy at the University of Saskatchewan. She works in the interdisciplinary field of socio-hydrology linking social science knowledge and methods with climatology and hydrology.

Geoffrey J. D. Hewings is professor and director emeritus of the Regional Economics Applications Laboratory at the University of Illinois. He earned his Ph.D. from the University of Washington, Seattle.

Jason R. Jurjevich is an associate professor in the Toulan School of Urban Studies and Planning and acting director of the Population Research Center (PRC) at Portland State University in Portland, Oregon, USA.

Ayoung Kim is visiting scholar of Agricultural Economics at Mississippi State University, USA.

Jacques Ledent was, until his recent retirement, professor of Demography in the Centre-Urbanisation, Culture et Société at the Institut National de la Recherche Scientifique (INRS) in Montréal, Québec (Canada).

Changfeng Li is an urban planning engineer in the Department of Informatics, China Academy of Urban Planning and Design.

William Lisowski applies quantitative modeling and analysis in a wide range of fields, most recently collaborating on research in demographics and population geography. His earlier work includes public policy analysis, information technology and telecommunications, and marketing.

Nan Liu is a research assistant in the School of Demography at the Australian National University.

David C. Maré is senior fellow at Motu Economic and Public Policy Research, Wellington, New Zealand.

Beth Mitchneck is a geographer who works on migration with a dual specialty on gender equity in academic STEM. She has held numerous faculty and administrative positions at the University of Arizona and University of Massachusetts at Lowell.

Wangshu Mu is a postdoctoral research associate in the School of Geographical Sciences and Urban Planning, Arizona State University.

Gordon F. Mulligan is professor emeritus of Geography and Development at the University of Arizona. He is a fellow of the Regional Science Association International and the Western Regional Science Association. His interests include economic geography, demography, spatial welfare, and urbanization.

K. Bruce Newbold is a professor of Geography in the School of Geography and Earth Sciences at McMaster University, Canada.

Helena A. K. Nilsson is a doctoral candidate in Economics at Jönköping International Business School, Jönköping, and the Institute of Retail Economics (HFI), Stockholm, where she worked as a research assistant prior to enrolling in the doctoral program. She has a bachelor and master's degree in Economics from the University of Gothenburg, School of Business, Economics and Law, Gothenburg.

Jacques Poot is visiting professor in the Department of Spatial Economics at Vrije Universiteit Amsterdam, the Netherlands, and emeritus professor at the National Institute of Demographic and Economic Analysis, University of Waikato, Hamilton, New Zealand.

James Raymer is a professor of Demography in the School of Demography at the Australian National University.

Peter A. Rogerson is at the University at Buffalo, where he is SUNY distinguished professor of Geography.

Daoqin Tong is an associate professor in the School of Geographical Sciences and Urban Planning at Arizona State University.

Brigitte S. Waldorf is professor of Agricultural Economics at Purdue University, USA.

Chapter 1
I Dream of Gini: Measures of Population Concentration and Their Application to US Population Distribution

Peter A. Rogerson

Abstract The unevenness of population across space may be captured with various measures of inequality; the Lorenz curve, the Gini coefficient, and the Hoover index constitute prime examples of such measures. In this chapter, I first review the history of these measures and then provide a selective review of their use in examining population concentration and deconcentration. Next, I show how the Gini coefficient may be disaggregated to show how population concentration varies within different ranges of population densities. The Gini coefficient is written as a weighted sum of the Hoover indexes for each population density category, where the weights are the proportion of total area in that density category. This disaggregation is applied to the US population for the period 2000–2015. Results show that population deconcentration is occurring among the subset of counties that have high population density and concentration is occurring among counties that have medium population density.

Keywords Gini coefficient · Hoover index · Population deconcentration

I am pleased to be able to contribute to this volume in honor of David Plane's contributions to the fields of geography and regional science. Dave and I first met in 1980 at the US Census Bureau, where we were doctoral students working on a project jointly funded by the American Statistical Association, the National Science Foundation, and the Census Bureau. Since that time, we have enjoyed both a career-long collaboration that has been productive and rewarding and a lifelong friendship that has been enjoyable and genuine. Over the course of our careers, we worked together on many topics at the intersection of demography and geography. Some of our early direction and inspiration came from Andrew Isserman, the leader of our project at the Bureau. Following Andy's passing in 2010, Dave and I coauthored a

P. A. Rogerson (✉)
Departments of Geography and Biostatistics, Wilkeson Hall, University at Buffalo, Buffalo, NY, USA
e-mail: rogerson@buffalo.edu

R. S. Franklin (ed.), *Population, Place, and Spatial Interaction*, New Frontiers in Regional Science: Asian Perspectives 40,
https://doi.org/10.1007/978-981-13-9231-3_1

paper on the Hoover index of population concentration for a special volume of the *International Regional Science Review* that was put together in memory of Andy. With this chapter, I am delighted to return to this topic. A special word about the title—over the years, Dave and I shared both hotel rooms at conferences and tents in campgrounds while on bike trips. In addition to the many stimulating conversations, Dave learned that I often have vivid dreams—these ranged from falling out of both my bed and then our 35th story window at the Bonaventure Hotel in Los Angeles to shouting away the intruders of my dreams that lurked outside of our tent. I think, then, it is only fitting to include an allusion to this in the title of the chapter.

1.1 Introduction

Substantial change has taken place in the distribution of population across counties over time in the United States. On a long time scale, the population has of course become urbanized, rising from 5.1% urban at the time of the first decennial census in 1790 to 80.7% urban in 2010. An increase has occurred in every decade with the exception of 1810–1820, when the urban population went from 7.3% to 7.2% (US Bureau of the Census 2012).

With respect to more recent change, a plethora of studies emerge whenever new census data are released. A Pew research report notes that urban counties have grown at about the nationwide rate of 13% during the period 2000–2016, while suburban and small urban counties have grown more rapidly (16%) (Parker et al. 2018). Simultaneously, rural counties have grown more slowly, rising just 3% over the period.

Similarly, Frey (2018), in a report from the Metropolitan Policy Program of the Brookings Institution, notes that urban core counties grew more slowly (at an annual rate of approximately 0.5%) than did mature and emerging suburban counties and exurban counties (annual rate of a bit over 1%) during the period 2000–2017. Toward the end of that period, small metropolitan areas began to grow at the expense of large metropolitan areas.

A widely recognized confounding factor in these studies is the definition of terms such as rural, suburban, urban, and exurban. Definitions vary from study to study, and the census definition itself has varied over time. Hall, Kaufman, and Ricketts (2006) provide an overview of many of the available definitions.

1.2 Some Measures of Concentration: With Historical Notes

One of the earliest measures of concentration was described by Lorenz in 1905; the well-known Lorenz curve provides a convenient way to visualize inequality. It is most commonly used for the depiction of income inequality. In that context, after arranging individuals (or groups of individuals) in terms of increasing income, cumulative population is plotted on the horizontal axis against cumulative income on the vertical axis. When there is little inequality, the plot will lie close to the 45° line. As inequality increases, the Lorenz curve bows further outward and away from (and below) the 45° line.

Shortly after, Gini (1912, 1914) introduced a numerical measure of inequality that is directly related to the Lorenz curve. In particular, the Gini coefficient is equal to the fraction of the area lying below the 45° line that is between the 45° line and the Lorenz curve. The Gini coefficient is also equal to half of the relative mean absolute difference between all pairs of incomes:

$$G = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} \sum \mid x_i - x_j \mid}{2n^2 \bar{x}}$$

The measure is variously known as the Gini coefficient (1,170,000), Gini index (871,000), and the Gini ratio (141,000), where the numbers in parentheses indicate the number of "hits" in Google (recognizing of course that this is not a perfect indicator of popularity or use).

As noted by Ceriani and Verme (2012), Gini (1912) first described his index in terms of mean differences, in a publication in Italian that has never been translated into English. In that publication, Gini discusses 13 different measures of inequality, of which the Gini coefficient is one. He draws an explicit connection to the Lorenz curve in his 1914 publication—this was eventually published in English, but not until 2005 (Gini 2005). Ceriani and Verme suggest that Gini's work became more widely known and disseminated following a note by Gini (1921) in English in *Economic Journal*, where he brought readers to the attention of the work of several Italian researchers. The note focused on measures of inequality and was written in response to an article on the topic in that journal by Dalton (1920).

The Hoover index of concentration, as it is known to geographers and demographers, is equal to the maximum vertical distance between the Lorenz curve and the 45° line. It is also equal to half of the relative mean absolute deviation (from the mean):

$$H = \frac{\sum \mid x_i - \bar{x} \mid}{2n\bar{x}}$$

In the context of income, it is interpreted as the percentage of total income that would have to be redistributed from the rich (defined as those to the right of the vertical line) to the poor (defined as those to the left of the vertical line), to equalize incomes. In the context of population, it is the percentage of population that would have to move from high-density places to low-density places to equalize population density across all spatial units. The Hoover index (13,000) is known variously as the Pietra index (3000), the Robin Hood index (15,000), and the Schutz index (47,000), where the number in parentheses is again the rounded number of "hits" in a Google search.

Hoover's 1936 paper is often cited as the one where he introduced what is now known as Hoover index of concentration. However, in that paper he used the Gini index, and there is no mention of his index of concentration. He cited Gini as an important contributor to measures of inequality and thanked the Nobel Prize economist Wassily Leontief for pointing this out to him, but he does not appear to have noticed at that time that Gini was responsible for the development of the coefficient; instead, he credited Vinci (1934) for interpreting the coefficient.

Hoover introduced his index of concentration in his 1941 paper. Here he credits Florence and Wensley (1939) for developing the measure, citing their book in the context of the extent of localization of manufacturing industries. Florence (1953) himself notes that he and Wensley first reported their initial work on this measure in 1937, in *Economic Journal*. And interestingly, in this 1937 work, they cite the work of Hoover as one of several researchers working on *other* measures.

Hoover's index actually appears to have first been discussed by Pietra (1915; translated in Pietra 2014), who showed the relations between several measures of inequality/concentration. Giorgi (2014) notes that Pietra showed the "relations existing between the mean difference, the simple mean deviations from the arithmetic mean and the median, and their geometrical interpretation". Schutz (1951) was a graduate student at Berkeley when he published his explication of the index (apparently without knowledge of and certainly without attribution of the earlier work of Pietra, Florence and Wensley, and Hoover). Despite the existence of earlier formulations, the term "Schutz index" is used much more often than the other names for the measure, perhaps due to its introduction and use within the field of economics.

1.3 Selected Review of Studies of Population Concentration and Deconcentration in the United States

The following review is meant to be illustrative and representative of the types of studies that have focused on population concentration and its measurement in the United States; it is by no means meant to be a comprehensive review.

Duncan et al. (1961) used the Hoover index to show how the US population became increasingly concentrated during the first half of the twentieth century. Lichter (1985) also used the Hoover index, focusing upon the latter half of the century, and the differing degrees of concentration by race.

Plane and Mulligan (1997) argue for the use of the Gini coefficient in population and migration research. They calculate and interpret several Gini coefficients in the context of the US migration system, concentrating on the measurement of the amount of spatial focusing that occurs—either within the entire system or within the sets of inflows and outflows. Rogers and Sweeney (1998) and Rogers and Raymer (1998) apply these measures and the coefficient of variation in their own analyses of US population redistribution.

Lichter and Johnson (2006) examine the spatial patterns of concentration and deconcentration among the foreign-born population that occurred during the 1990s. They find that (a) the foreign-born are dispersing away from metropolitan, gateway cities (although they remain much more concentrated than the native-born) and (b) they are less segregated from other populations than they were in the past—so-called balkanization and isolation are not as acute as they once were.

Plane et al. (2005) examine population distribution in the context of the urban hierarchy. They place their work in historical context by noting that over the long-term net migration has been primarily up the urban hierarchy, leading to increasing population concentration. They use migration data for the period 1995–2000 to look at net flows up and down the hierarchy—the nature of such flows can differ substantially according to age and life course stages and events (such as college, military service, family formation, retirement, etc.).

By far, the majority of attention to population concentration and deconcentration in the United States has been devoted to the "rural renaissance" of the 1970s and the subsequent relative growth rates of rural, urban, and suburban areas. Vining and Strauss (1977) argued that the recently observed deconcentration was a "clean break" from past trends. At about the same time, McCarthy and Morrison (1977) also noted the increased net in-migration that was being experienced by rural, nonmetropolitan counties, signaling the beginning of a new or at least more complex demographic trend. They also found that retirement and recreation were increasingly important as drivers of this migration.

Gordon (1979) argued that the newfound reversal was perhaps due to growth that was extending from metro areas and spilling over into adjacent nonmetropolitan counties and that the "clean break" hypothesis deserved further scrutiny. He used the Hoover index to support this view using data from 18 countries. Bourne (1980) attributed the decreasing concentration to implicit US policies that favored the growth of exurbia (and the lack of policies that encouraged redevelopment within cities).

Work in this area developed rapidly, with Berry (1980) providing an early review of the counterurbanization trend and John Long (1981) producing a book describing the trends toward deconcentration at various spatial scales. Lichter and Fuguitt (1982) focused some of their attention on population distribution *within* nonmetropolitan areas, showing that there was deconcentration at that level as well. Like Morrison and McCarthy, they found that economic explanations for demographic change in nonmetropolitan areas were of decreasing importance.

The empirical work also led to the development of various theoretical frameworks within which the changes could be understood more broadly. Morrill (1979, 1980) and Geyer and Kontuly (1993) developed conceptual and theoretical frameworks for population and concentration; the latter authors, for example, use data from different countries to suggest that counterurbanization represents the final phase of the first cycle of demographic change in an urban system, and this is followed by a cycle where concentration is again dominant. Morrill (1980) argued against a "clean break" in the 1970s, suggesting that long-term agglomerative forces acted differentially across space and that older and denser places experience out-migration and deconcentration as a stage in the evolution of the geographic landscape.

Fuguitt (1985) was an early reporter of the reversal of the turnaround, reporting that there was a return to concentration in the metropolitan nonmetro system during the early 1980s. Fuguitt also reviewed the literature on the turnaround, emphasizing the point that the focus on trends in concentration and deconcentration had the beneficial effect of drawing more attention to other facets of migration research, including individual migration behavior, preferences, and the relation of migration to employment. Finally, he pointed out that the research generated by the topic was not only broad in scope but also large in its volume. At that time Fuguitt speculated that a return to concentration seemed unlikely. Cochrane and Vining (1988) also provided early evidence of the reversal of counterurbanization that occurred during the 1980s.

Then, during the 1990s, there was a return to the deconcentration witnessed in the 1970s. Long and Nucci (1997a) documented this return to deconcentration using the Hoover index, with counties as the spatial unit. In a second paper, Long and Nucci (1997b) documented this further, simultaneously (a) correcting an error in the original county-based series of Hoover indexes reported by Duncan et al. (1961), (b) extending the Duncan et al. time series both backward and forward in time, and (c) noting that there was also deconcentration at the county level between 1890 and 1910. Johnson and Cromartie (2006) provided corroborating evidence for the renewed deconcentration that took place during the 1990s by using data from the 2000 census.

Domina (2006) found that trends changed again during the late 1990s and the early years of the twenty-first century, with nonmetro areas once again experiencing net out-migration. Domina used data from the Current Population Survey to show that much of the net out-migration was attributable to those with higher levels of education and to suggest that economic factors now carry important explanatory power in understanding nonmetro migration.

Rogerson and Plane (2013) calculated annual Hoover indexes for the period 1990–2009 at a variety of spatial scales, including the county level. Like Domina, they find concentration to be generally increasing at the county level from the late 1990s. Interestingly, they find that births, deaths, and immigration all caused the index of concentration to increase, but net internal migration on its own would have led to deconcentration during the period.

1.4 Measurement: Disaggregating the Lorenz Curve

When the Lorenz curve is split into two pieces by the vertical line representing the Hoover index and the maximum difference between cumulative populations and cumulative areas, each of the two pieces may then be scaled and transformed into a Lorenz curve. Each of these has an associated Hoover index. If desired, each of these in turn may also be split into two pieces and Hoover indexes calculated for each of the new pieces; this process can continue to be repeated. Thus it is possible to examine inequality along portions of the Lorenz curve—in the application that is the focus here, we may examine population concentration for sets of regions that have a particular range of population density.

The Gini coefficient is always at least as high as the Hoover index. The excess is equal to twice the area between the Lorenz curve and the 45° line that lies beneath the triangle created by the end points of the Lorenz curve and the lower point of the vertical line associated with the Hoover index. The Gini coefficient for the full Lorenz curve may be calculated as the sum of the overall Hoover index and a weighted sum of the Hoover indexes associated with the disaggregation described above, where the weights are equal to the proportion of area (or population, in the case of application to income inequality) associated with that segment of the curve (Rogerson 2013).

We now begin with an illustration for the US population in 2010, at the county level. The Hoover index was 66.16 (after rounding to two decimal places); 16.31% of the population lived on 82.48% of the land (and, of course, the other 83.69% of the population lived on just 17.52% of the land). The former group consisted of 2166 low-density counties, and the latter group had 977 high-density counties. The low-density counties had an average of 17.29 people per square mile; the high-density counties averaged 417.5 people per square mile.

Each of these two groups may be examined as a separate subsystem. The line 0B in triangle 0AB in Fig. 1.1 represents what would be expected if all low-density counties had the same density. Similarly, the line BD in triangle BCD represents what would be expected if all high-density counties had the same population density. There is of course concentration *within* low-density and high-density subsystems, and this is captured by these two triangles.

Among the 2166 low-density counties, the 795 with the lowest density have 15.96% of the population on 65.73% of the land area ($H_L = 65.73 - 15.96 = 49.77$). Among the 977 high-density counties, 687 have 34.29% of the population on 76.95% of the area, for a Hoover index of $H_U = 76.95 - 34.29 = 42.66$.

These three breakpoints—one for the entire Unites States and the other two for low- and high-density subsystems—serve to divide the United States into four categories of population density (labeled 1 through 4 on Fig. 1.1). The coordinates of points A, B, C, and D, representing cumulative population and cumulative area, are given in Table 1.1. The table also reports the absolute values of the differences between cumulative area and cumulative population. Note that the highest of these absolute differences (when multiplied by 100) represents the greatest vertical distance between the Lorenz curve and the 45° line and is equal to the Hoover index of 66.16.

The triangles for the low-density and high-density subsystems are redrawn in Figs. 1.2 and 1.3, respectively. Part (a) of each figure is the triangle as it appears within the original Lorenz curve in Fig. 1.1; part (b) of each figure has scaled the triangle in part (a) to represent a Lorenz curve for the two-region subsystem that the figure represents.

The length of the maximum vertical distance in part (a) of Figs. 1.2 and 1.3 can be thought of scaled versions of the full Hoover indexes found in part (b). For Fig. 1.2a, the y-coordinate of point P is found as (0.1631/0.8248) (0.5422) = 0.1072. The vertical distance and scaled Hoover index is then $0.1072 - 0.0260 = 0.0812$. In Fig. 1.3a, point P has a y-coordinate of $0.1631 + (0.9596 - 0.8248)(1 - 0.1631)/(1 - 0.8248) = 0.8070$. The vertical distance and scaled Hoover index is $0.8070 - 0.4501 = 0.3569$.

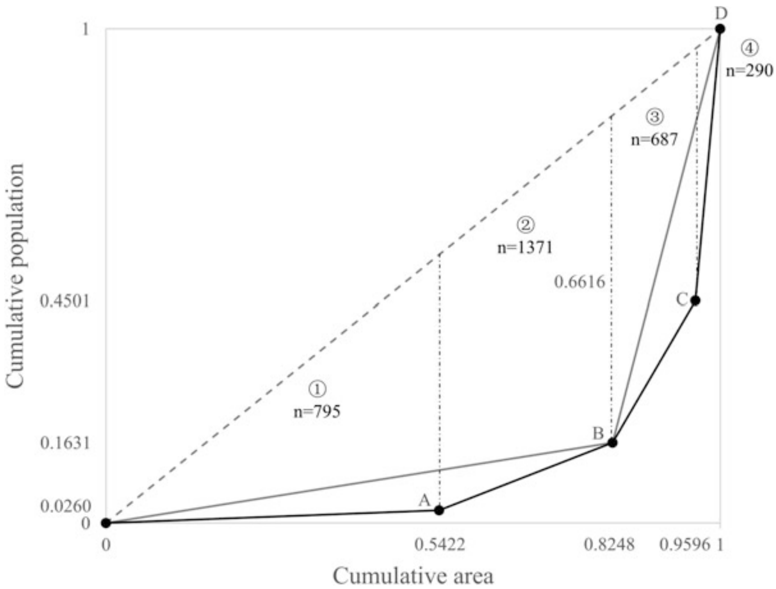


Fig. 1.1 Lorenz curve for US population: 2010

Table 1.1 Areas and populations for US counties, 2010

Population density category	Proportion of total population (p)	Proportion of total area (a)	Cumulative population (cp)	Cumulative area (ca)	Absolute difference ($cp - ca$)
1	0.0260	0.5422	0.026	0.5422	0.5161
2	0.1371	0.2824	0.1631	0.8248	0.6616
3	0.2970	0.1348	0.4501	0.9596	0.5095
4	0.5499	0.0404	1.0	1.0	0

The Gini coefficient (for the Lorenz curve in Fig. 1.1) may be calculated as a weighted sum of these scaled Hoover indexes, where the weights are the proportion of total land area that is associated with each section of the Lorenz curve. Thus

$$G = .6616 + (.8248)(.0812) + (.1752)(.3569) = .7911.$$

The second and third terms in the sum also represent the areas of triangles OAB and BCD in Fig. 1.1 (repeated as Figs. 1.2a and 1.3a), respectively.

More generally, suppose p_i is the proportion of total population found in subset i, and a_i is the proportion of total area found in subset i (where i is indexed from 1 to 4, corresponding to the sections in Fig. 1.1). Let cp_i and ca_i be the corresponding cumulative proportions. Furthermore, subsets 1 and 2 together comprise the low-density population and subsets 3 and 4 together comprise the high-density population (as determined by the system-wide Hoover index). The Hoover index for the entire system is equal to the absolute value of the difference between cp_2 and ca_2.

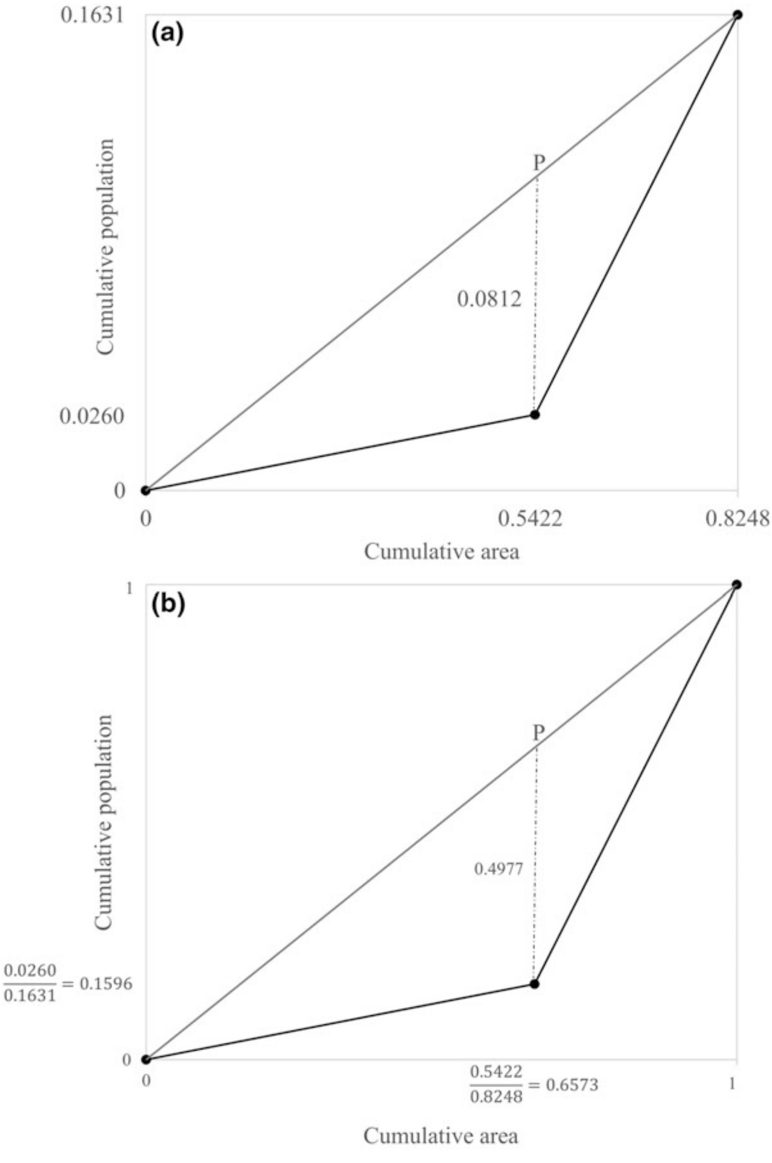


Fig. 1.2 Lorenz curve for low density US counties: 2010

The Hoover index for low-density subsystem in Fig. 1.2b (say H_L) is the absolute value of the difference between cp_1/cp_2 and ca_1/ca_2. The Hoover index for high-density subsystem in Fig. 1.3b, comprised of Sects. 3 and 4 in Fig. 1.1, is the absolute value of the difference between $(ca_3 - ca_2)/(1 - ca_2)$ and $(cp_3 - cp_2)/(1 - cp_2)$.

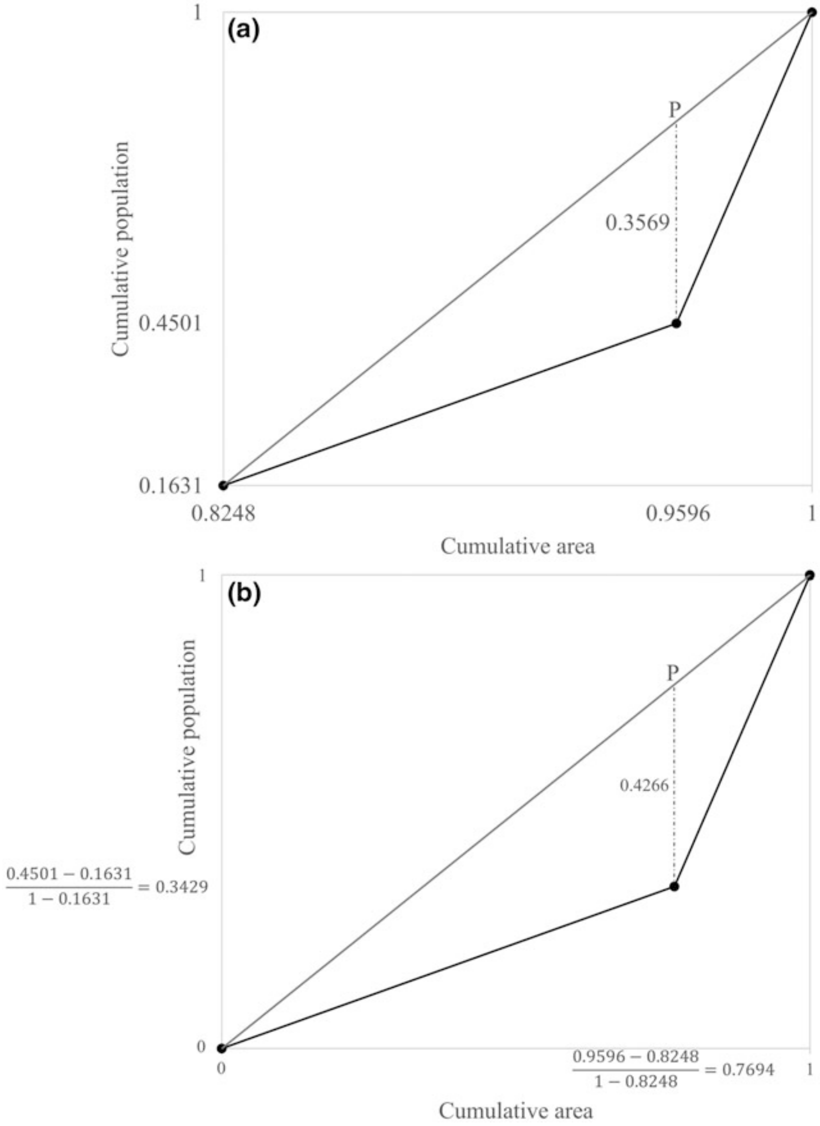


Fig. 1.3 Lorenz curve for high density US counties: 2010

The scaled Hoover indexes (H′), i.e., the heights of the two sub-triangles in Fig. 1.1, are found as the difference between the y-coordinate of point P and cp_1 in Fig. 1.2a and the difference between the y-coordinate of point P and cp_3 in Fig. 1.3a. The y-coordinate of point P in Fig. 1.2a is equal to $(cp_2/ca_2)(ca_1)$; in Fig. 1.3a it is equal to $cp_2 + (ca_3 - ca_2)(1 - cp_1)/(1 - ca_1)$. Thus

$$H'_L = (cp_2/ca_2)(ca_1) - cp_1$$
$$H'_U = cp_2 + (ca_3 - ca_2)(1 - cp_1)/(1 - ca_1) - cp_3$$

and this last equation can be simplified a bit by replacing the difference in cumulative proportions ($ca_3 - ca_2$) by the proportion a_3 and ($cp_2 - cp_3$) by $-p_3$:

$$\begin{aligned} H'_U &= cp_2 + (a_3)(1 - cp_1)/(1 - ca_1) - cp_3 \\ &= (a_3)(1 - cp_1)/(1 - ca_1) - p_3 \end{aligned}$$

The Gini coefficient is then equal to H + (ca_2) H_L' + (1 − ca_2) H_U'.

1.5 Changes in Population Concentration: 2000–2015

In 2000, when the 2144 counties with lowest pop density are examined (out of a total of 3141 counties), we find that they contain 82.3% of the country's area, but just 16.7% of the population, leading to a Hoover index of 65.61. These counties have an average population density of 16.14 people per square mile; the remaining 997 high-density counties have an average population density of 374.81 people per square mile. During the first decade of this century then, there was an increase in population concentration at the county level, as the Hoover index rose from 65.61 to 66.16. By 2015, the US population was even more concentrated at the county level, with H = 66.73.

Among the 2144 low-density counties, the Hoover index was 49.31 in 2000, with the most rural of places (758 of the 2144 counties) containing 65.53% of the land area in this subset but just 16.22% of the population and having an average density of 3.99 people per square mile. The remaining 2144 − 758 = 1386 counties had an average population density of 39.22 people per square mile. As a group, the low-density counties experienced a small increase in concentration, with the Hoover index rising from 49.31 to 49.77 during the decade. Concentration among the low-density counties continued during the first half of the next decade, with the Hoover index rising further to 49.87 by 2015.

Among the 997 high-density counties, the 719 least dense of these counties had 78.06% of the area but just 33.98% of the population (and an average population density of 163.04 people per square mile), leading to a Hoover index of 44.08 in 2000. The remaining high-density counties had an average population density of 1131.2 people per square mile. Within the high-density counties, deconcentration was experienced during the decade, with the Hoover index falling from 44.08 to 42.66. From 2010 to 2015, the index then rose slightly, to H = 42.7. These results are summarized in Table 1.2.

Table 1.2 Summary of changes in population concentration: 2000–2015

	2000	2010	2015
n counties	3141	3143	3143
Total pop	281,421,906	308,745,538	321,396,328
Total area	3,537,438	3,531,905	3,531,905
Pop density	79.56	87.42	91
H	65.61	66.16	66.73
Low-density counties			
(*n*)	2144	2166	2185
Area	2,911,440 (0.823)	2,912,978 (0.8248)	2,923,256 (0.8277)
Population	46,977,694 (0.167)	50,370,014 (0.1631)	51,530,769 (0.1603)
Density	16.14	17.29	17.63
High-density counties			
(*n*)	997	977	958
Area	625,498	618,927	608,649
Population	234,444,212	258,375,524	269,865,559
Density	374.81	417.46	443.38
Low-density counties *H*	49.31	49.77	49.87
Lower density portion of low density counties			
(*n*)	758	795	816
Area	1,907,779 (0.6553)	1,914,827 (0.6573)	1,927,283 (0.6593)
Population	7,618,610 (0.1622)	8,038,937 (0.1596)	8,276,116 (0.1606)
Density	3.99	4.2	4.29
Higher density portion of low-density counties			
(*n*)	1386	1371	1369
Area	1,003,661	998,151	995,973
Population	39,359,084	42,331,077	43,254,653
Density	39.22	42.41	43.43
High-density counties *H*	44.08	42.66	42.72
Lower density portion of high-density counties			
(*n*)	719	687	672
Area	488,672 (0.7806)	476,267 (0.7695)	457,974 (0.7524)
Population	79,673,800 (0.3398)	88,603,570 (0.3429)	87,775,140 (0.3253)
Density	163.04	186.04	191.66
Higher density portion of High-Density Counties			
(*n*)	278	290	286
Area	136,826	142,660	150,675
Population	154,770,412	169,771,954	182,090,419
Density	1131.15	1190.05	1208.5

Note: Areas are in square miles; densities are people per square mile

Table 1.3 Hoover index along four segments of Lorenz curve: 2000 and 2010

Year	2000	2010	2015
Low density portion of low-density counties			
H	43.42	43.43	43.68
% population	20.58	20.03	21.62
% area	64	63.45	65.3
n	286/758	310/795	330/816
High density portion of low-density counties			
H	17.93	18.65	18.95
% population	39.25	38.14	37.84
% area	57.18	56.78	56.79
n	730/1386	720/1371	722/1369
Low density portion of high-density counties			
H	20.25	20.95	20.7
% population	39.3	39.44	40.45
% area	59.54	60.39	61.15
n	425/719	412/687	399/672
High density portion of high-density counties			
H	30.98	29.57	29.96
% population	36.48	36.94	37.29
% area	67.45	66.51	67.24
n	160/278	169/290	161/286

Table 1.3 shows the results of further disaggregation, where there are four separate Hoover indexes associated with eight sections of the Lorenz curve. There has been little change in concentration over the period for the subset of lowest density counties. For the next level up the density hierarchy, there has been a steady increase in concentration between 2000 and 2015, with H increasing from 17.93 to 18.95 (and it is interesting to note that this subsystem of counties has the lowest H values, indicating relative uniformity in density). At the next step up the curve, there was also an increase in concentration between 2000 and 2010 and then a slight decrease during the first half of this decade. Finally, for the subset of counties with the highest densities, there was deconcentration during 2000–2010, followed by a slight increase in concentration during 2010–2015.

1.6 Attributing Change to Particular Counties

The Hoover index does not change when people move from a region on one side of the vertical line that defines the Hoover index to another region on that same side of the line. Migration will only cause change in the index when people move from a

Table 1.4 Counties contributing to increasing concentration within the high density portion of the low density counties

State	County	Population 2000	Population 2015	Percent change, 2000–2015
Arizona	Pinal County	179,277	406,584	126.8
Colorado	Mesa County	116,255	148,513	27.7
Colorado	Summit County	23,548	30,257	28.5
Florida	Baker County	22,259	27,420	23.2
Florida	Wakulla County	22,863	31,535	37.9
Georgia	Bacon County	10,103	11,299	11.8
Georgia	Candler County	9577	10,886	13.7
Georgia	Lanier County	7241	10,312	42.4
Georgia	Long County	10,304	17,731	72.1
Georgia	Pulaski County	9588	11,396	18.9
Hawaii	Hawaii County	148,677	196,428	32.1
Minnesota	Mille Lacs County	22,330	25,788	15.5
Montana	Missoula County	95,802	114,181	19.2
Oklahoma	Marshall County	13,184	16,232	23.1
Oregon	Deschutes County	115,367	175,268	51.9
Oregon	Hood River County	20,411	23,137	13.3
Texas	Austin County	23,590	29,563	25.3
Texas	Burnet County	34,147	45,463	33.1
Texas	Kendall County	23,743	40,384	70.0
Texas	Polk County	41,113	46,972	14.3
Texas	San Jacinto County	22,246	27,413	23.2
Texas	Somervell County	6809	8739	28.3
Utah	Washington County	30,373	34,765	14.4

region on one side of the vertical line to a region on the other side of the line. When two Lorenz curves are compared, there may be regions that are on one side of the vertical line at one point in time and on the other side of the line at the next point in time. A region on the left side of the vertical line in 1 year will contribute to increased concentration if it is on the right side of the line in a subsequent year.

To illustrate, here we examine the set of counties that had a density of about 40 persons per square mile in 2000 and find those that contributed to the increased concentration that was observed over the next 15 years. In 2000, there were 1386 counties in Sect. 2 of the Lorenz curve (when it is divided into four sections, in order of increasing density); there were 1369 counties in that section in 2015. In 2000, these counties had a population density of 39.22 persons per square mile, and they ranged in population density from 16.2 to 79.4 people per square mile. By 2015, this section of the curve contained counties with an average population density of 43.43 persons per square mile and densities ranging from 17.7 to 90.8 persons per square mile. The 23 counties listed in Table 1.4 were among the low-density counties in this portion of the Lorenz curve in 2000 but were in the high density portion in 2015. They therefore contributed to the increased concentration in this subsystem. Many of these grew rapidly and are easily recognizable as amenity-laden places. Others, such as Bacon County, Georgia, did not grow particularly fast but were close to the vertical line in the Lorenz curve of 2000 (near the mean density of 39.22 persons per square mile); even a small amount of growth was sufficient to put them on the right-hand side of the vertical line in the 2015 Lorenz curve.

1.7 Discussion and Summary

The Hoover index of concentration has a long history of application in population geography and demography. It has received particularly widespread use in the study of rural population change since the 1970s.

Like the Hoover index, the Gini coefficient may be interpreted in relation to the Lorenz curve. It may also be derived as the sum of the Hoover index, plus weighted, scaled Hoover indexes associated with partitions of the Lorenz curve constructed by dividing the curve at the vertical line associated with the Hoover index. This approach was used here to show how population concentration has changed during the period 2000–2015. Overall increases in concentration were accompanied by concentration for many population density levels, but deconcentration occurred for that set of counties comprising the highest density levels.

It would be interesting to assess the temporal change in Hoover indexes for these subsystems for various demographic components of change. Rogerson and Plane (2013) found increasing concentration for births, deaths, and international migration and *deconcentration* for internal migration. Whether these trends hold at all levels of population density is unknown. Rural areas, e.g., tend to have a relatively older age structure, and the relatively large number of deaths and small number of births act to increase population concentration in the more dense places among them. This effect might, for example, be less apparent or nonexistent at higher population densities. It would be particularly interesting to know, e.g., whether net internal migration is leading to deconcentration at all levels of population density or whether only a certain portion of the density hierarchy is witness to that trend.

Finally, it could also be enlightening to examine more closely the geographic distribution of these changes. The focus here has been on groupings of counties by population density. These groupings could easily be mapped, as could the counties that increased or decreased in population density sufficiently to contribute to a change in the Hoover index. Similarly, the analysis described above could be carried out for regional divisions of the United States.

Acknowledgments I would like to thank Xiang Ye for his assistance with the figures.

References

Berry BJL (1980) Urbanization and counterurbanization in the United States. Ann Am Acad Pol Soc Sci 451(1):13–20
Bourne LS (1980) Alternative perspectives on urban decline and population deconcentration. Urban Geogr 1:39–52
Ceriani L, Verme P (2012) The origins of the Gini index: extracts from Variabilitia e Mutabilita (1912) by Corrado Gini. J Econ Inequal 10:421–443
Cochrane SG, Vining DR Jr (1988) Recent trends in migration between core and peripheral regions in developed and advanced developed countries. Int Reg Sci Rev 11:215–243
Dalton H (1920) The measurement of the inequality of incomes. Econ J 30:348–461
Domina T (2006) What clean break? Education and nonmetropolitan migration patterns, 1989–2004. Rural Sociol 71(3):373–398
Duncan OD, Cuzzort R, Duncan B (1961) Statistical geography. Free Press, Glencoe
Florence PS (1953) The logic of British and American industry: a realistic analysis of economic structure and government. Routledge and Kegan Paul, London
Florence PS, Wensley AJ (1939) The location of industry in Great Britain. London
Frey WH (March 26, 2018) US population disperses to suburbs, exurbs, rural areas, and “middle of the country” metros. https://www.brookings.edu/blog/the-avenue/2018/03/26/us-population-disperses-to-suburbs-exurbs-rural-areas-and-middle-of-the-country-metros/. Last accessed 2 Dec 2018
Fuguitt G (1985) The nonmetropolitan population turnaround. Annu Rev Sociol 11:259–280
Geyer HS, Kontuly T (1993) A theoretical foundation for the concept of differential urbanization. Int Reg Sci Rev 15(2):157–177
Gini C (1912) Variabilit’a e Mutabilit’a. Contributo allo studio delle distribuzioni e delle relazioni statistiche, Studi economico-giuridici Anno III, Parte II. Facolt’a di giurisprudenza della Regia Universit’a di Cagliari. Cuppini, Bologna
Gini C (1914) Sulla Misura della Concentrazione e della Variabilit’a dei Caratteri. Atti del Reale Istituto Veneto di Scienze, Lettere ed Arti LXXIII(II):1203–1248
Gini C (1921) Measurement of inequality of incomes. Econ J 31:124–126
Gini C (2005) On the measurement of concentration and variability of characters. METRON - Int J Stat LXIII(1):3–38
Giorgi GM (2014) A couple of good reasons to translate papers of the Italian statistical tradition. Metro 72(1):1–3
Gordon P (1979) Deconcentration without a ‘clean break’. Environ Plan A 11(3):281–289
Hall SA, Kaufman JS, Ricketts TC (2006) Defining urban and rural areas in U.S. epidemiological studies. J Urban Health 83(2):162–175
Hoover EM (1936) The measurement of industrial localization. Rev Econ Stat 18:162–171
Hoover E (1941) Interstate redistribution of population: 1850–1940. J Econ Hist 1:199–205
Johnson KM, Cromartie JB (2006) The rural rebound and its aftermath. In: Kandel WA, Brown DL (eds) Population change and rural society. Springer, Dordrecht, pp 25–49
Lichter DT (1985) Racial concentration and segregation across US counties, 1950–1980. Demography 22(4):603–609
Lichter DT, Fuguitt GV (1982) The transition to nonmetropolitan population deconcentration. Demography 19(2):211–221
Lichter DT, Johnson KM (2006) Emerging rural settlement patterns and the geographic redistribution of America’s new immigrants. Rural Sociol 71(1):109–131
Long J (1981) Population deconcentration in the United States. Department of Commerce, US Government Printing Office, Washington, DC
Long L, Nucci A (1997a) The ‘clean break’ revisited: is US population again deconcentrating? Environ Plan A 29(8):1355–1366
Long L, Nucci A (1997b) The Hoover index of population concentration: a correction and update. Prof Geogr 49(4):431–440

Lorenz MO (1905) Methods of measuring the concentration of wealth. Q Publ Am Stat Assoc ix:209–20I

McCarthy KF, Morrison PA (1977) The changing demographic and economic structure of nonmetropolitan areas in the United States. Int Reg Sci Rev 2(2):123–142

Morrill R (1979) Stages in patterns of population concentration and dispersion. Prof Geogr 31 (1):55–65

Morrill R (1980) The spread of change in metropolitan and nonmetropolitan growth in the United States, 1940–1976. Urban Geogr 1(2):118–129

Parker K, Horowitz JM, Brown A, Fry R, Chon D, Igielnik R (May 22, 2018) What unites and divides urban, suburban, and rural communities. Pew Research Report. http://www.pewsocialtrends.org/2018/05/22/demographic-and-economic-trends-in-urban-suburban-and-rural-communities/ Last accessed 2 Dec 2018

Pietra G (1915) Delle relazioni tra gli indici di variabilità (Nota I)*" Atti del Reale Istituto Veneto di Scienze, Lettere e Arti. 1915, Vol. LXXIV, Part I, pages 775–792*

Pietra G (2014) On the relationships between variability indices (note I) (translated from the 1915 article) in *Metron* 72(1):5–16

Plane DA, Mulligan GF (1997) Measuring spatial focusing in a migration system. Demography 34 (2):251–262

Plane DA, Henrie CJ, Perry MJ (2005) Migration up and down the urban hierarchy and across the life course. Proc Natl Acad Sci 103(43):15313–15318

Rogers A, Raymer J (1998) The spatial focus of US interstate migration flows. Int J Popul Geogr 4:63–80

Rogers A, Sweeney S (1998) Measuring the spatial focus of migration patterns. Prof Geogr 50 (2):232–242

Rogerson P (2013) The Gini coefficient of inequality: a new interpretation. Lett Spat Resour Sci 6 (3):109–120. https://doi.org/10.1007/s12076-013-0091-x

Rogerson PA, Plane DA (2013) The Hoover index of population concentration and the demographic components of change: an article in memory of Andrew M. Isserman. Int Reg Sci Rev 36(1):97–114

Schutz R (1951) On the measurement of income inequality. Am Econ Rev 41:107–122

US Bureau of the Census (2012) United States summary: 2010. 2010 census of population and housing, population and housing unit counts, CPH-2-5. U.S. Government Printing Office, Washington, DC

Vinci F (1934) Manuale di Statistica, I (Bologna), pp 113–118

Vining D, Strauss A (1977) A demonstration that the current deconcentration of population in the United States is a clean break with the past. Environ Plan A 9:751–758

Chapter 2
Unraveling David Plane's Tools for Analyzing the Income Impacts of Interregional Migration Flows

Jacques Ledent

Abstract Two decades ago, David Plane introduced several tools for analyzing the redistribution of money associated with interregional migration flows (Plane D, Int J Popul Geogr 5:195–212, 1999). But, although these seminal tools were totally correct, they did not deliver all of what they could aspire to because, in several instances, they were not pushed to their logical end. The purpose of this chapter thus is to revisit Plane's tools and to rework them to reveal their full potential for understanding how interregional migration flows affect a region's aggregate and per capita income.

Keywords Economic impacts · Income · Interregional migration · Interregional redistribution · Population · Regional modeling · Spatial analysis

2.1 Introduction

At the turn of millennium, upon observing that the literature on population migration had focused on its determinants at the expense of its impacts, David Plane embarked on a bold effort aimed at developing tools for assessing how interregional migration flows affect the redistribution of money in a system of regions, or, in short, for analyzing income migration. This effort led him to publish a pioneering paper (Plane 1999) in which he introduced three tools capable of illuminating how changes in income accrue to regions as a consequence of people moving between them. True, these tools are totally correct, but it remains, as will become clear, that they are not definitive and can be manipulated to improve our understanding of a key

This research was originally presented at the 20th conference of the Pacific Regional Science Conference Organization (PRSCO) held in Vancouver, BC, on May 6–10, 2007.

J. Ledent (✉)
Centre - Urbanisation Culture Société, Institut National de la Recherche Scientifique, Montréal, QC, Canada
e-mail: jacques.ledent@ucs.inrs.ca

R. S. Franklin (ed.), *Population, Place, and Spatial Interaction*, New Frontiers in Regional Science: Asian Perspectives 40,
https://doi.org/10.1007/978-981-13-9231-3_2

consequence of population migration: its influence on income migration. The purpose of this chapter thus is to revisit Plane's tools on income migration and to rework them to enhance their usefulness. This is accomplished by resorting to the very dataset used by Plane—that is, a dataset of 1993–1994 income migration within the system consisting of the 50 US states plus Washington, DC, referred to as the "US states."

The chapter consists of three main sections, one for each of the three tools introduced by Plane. Section 2.2 elaborates further Plane's decomposition of a region's change in aggregate income attributable to interregional migration flows into two components reflecting net population migration for one and the per capita income differential between in- and outmigrants for the other. Section 2.3 then zeroes in on Plane's income version of the index of population migration effectiveness commonly employed in migration research and comes up with an innovative decomposition of his index of income migration effectiveness into two components also reflecting net population migration and the per capita income differential between in- and outmigrants. Finally, Sect. 2.4 turns to Plane's decomposition of a region's change in per capita income into three elements that stress the differentials in per capita income between the three pools of population involved: inmigrants, outmigrants, and stayers. It eventually reshapes this decomposition in such a way that it consists of only two elements, which can be simply characterized as an inmigrant- and an outmigrant-related variable.

2.2 The Decomposition of a Region's Change in Aggregate Income into Its Constituent Population and Income Components

The first tool introduced by Plane in his pioneering paper is a technique for decomposing a region's change in aggregate income. It seeks to untangle the impacts of (1) net population migration and (2) the per capita income differential between in- and outmigrants on the interregional distribution of money as a result of people moving between regions. Let Y_I be the aggregate income of immigrants (inmigrant income):

$$Y_I = iI$$

where I is the flow of inmigrants (inmigrant flow) and i the per capita income of inmigrants. Similarly, let Y_O be the aggregate income of outmigrants (outmigrant income):

$$Y_O = oO$$

where O is the flow of outmigrants (outmigrant flow) and o the per capita income of outmigrants. Thus, net income migration, the difference between the aggregate incomes of immigrants and outmigrants, is

$$Y_N = Y_I - Y_O = iI - oO.$$

Basically, Plane decomposes this variable into two components:

- A net population migration component Y_N^{NMIG} which he defines as the product of the net migrant flow $(I - O)$ by a per capita income measure which he sets equal to the arithmetic average of the per capita incomes of in- and outmigrants $\frac{i+o}{2}$, expressed as

$$Y_N^{NMIG} = \frac{i+o}{2}(I - O).$$

- A differential income component $Y_N^{\Delta INC}$ which, on subtracting Y_N^{NMIG} from Y_N, proves to be the product of the arithmetic average of the in- and outmigrant flows $\frac{I+O}{2}$ by the difference between the per capita incomes of in- and outmigrants $(i - o)$, expressed as

$$Y_N^{\Delta INC} = iI - oO - \frac{i+o}{2}(I - O) = (i - o)\frac{I+O}{2}.$$

But, it is possible to go further than Plane. In the following expression obtained by adding the two components just distinguished

$$Y_N = \frac{i+o}{2}(I - O) + (i - o)\frac{I+O}{2},$$

factor the arithmetic average of the in- and outmigrant flows as well as the arithmetic average of the per capita incomes of in- and outmigrants. As a result, net income migration may be rewritten as

$$Y_N = 2\frac{i+o}{2}\frac{I+O}{2}\left[\frac{I-O}{I+O} + \frac{i-o}{i+o}\right]$$

in which $\frac{I-O}{I+O}$ is the classical index of population migration effectiveness, here denoted E, and $\frac{i-o}{i+o}$ is an index of the relative levels of per capita income among in- and outmigrants, here denoted e. Thus,

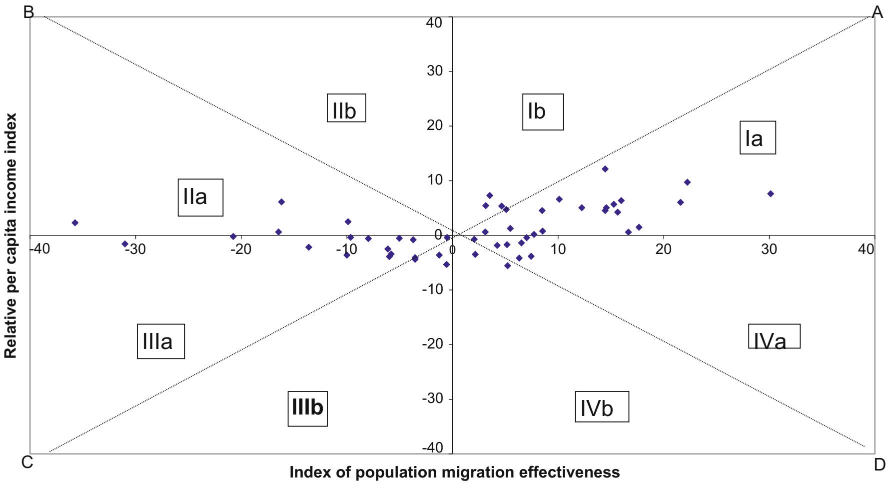


Fig. 2.1 Decomposition of net income migration into its net population migration and differential income components, US states, 1993–1994

$$Y_N = 2\frac{i+o}{2}\frac{I+O}{2}[E+e].$$

In other words, net income migration is equal to the sum of two indices, the index E of population migration effectiveness and the relative per capita income index e, up to a multiplicative factor that is equal to twice the product of the arithmetic average of the in- and outmigrant flows by the arithmetic average of the per capita incomes of in- and outmigrants. Moreover, the net (population) migration component Y_N^{NMIG} is proportional to E and the differential income component $Y_N^{\Delta INC}$ to e, up to the same multiplicative factor:

$$Y_N^{NMIG} = 2\frac{i+o}{2}\frac{I+O}{2}E$$

and

$$Y_N^{\Delta INC} = 2\frac{i+o}{2}\frac{I+O}{2}e.$$

Based on this, Fig. 2.1 offers a graphical representation of the decomposition of net income migration for the US states, one that supersedes the one provided in Plane's Fig. 2.3 in that it enables one to display the 51 data points. This was obtained by a slight modification of the axes. Instead of referring to the net population migration component Y_N^{NMIG} and the differential income component $Y_N^{\Delta INC}$, the axes now refer to the E and e indices, respectively, to which the two components of net income migration are equal, up to a common multiplicative factor.

Note that, as a consequence of their respective definition, the E and e indices take on values that are constrained to be between –1 and +1, or rather between −100 and +100 after multiplication of each index by 100.[1] But, with reference to the particular dataset at hand, the 51 data points are located within the ABCD square (for which the E and e indices lie between −40 and +40). Moreover, since the states having positive net income migration are such that $E + e$ is positive, the corresponding data points are necessarily located above the northwest to southeast (or BD) diagonal of the ABCD square. And conversely, the data points pertaining to those states having negative net income are located below that same diagonal.

Now, let the quadrants defined by the two axes be labeled as I to IV when moving counterclockwise from the quadrant with positive values of both E and e (quadrant I) toward the quadrant with positive values of E and negative values of e (quadrant IV). Moreover, upon joining in each of the quadrants I–IV the origin to the opposite corner of the quadrant (A, B, C, and D, respectively), thus resulting into eight triangles, let the letter a be assigned to the four triangles, a side of which goes along the horizontal axis, and the letter b to the other four triangles, a side of which goes, rather, along the vertical axis. In other words, the states located in the a-triangles are such that the absolute value of E (or the net population migration component) is larger than the absolute value of e (or the differential income component). Conversely, the states located in the b-triangles are such that the absolute value of E (or the net population migration component) is smaller than the absolute value of e (or the differential income component).

Thus, the label given to each triangle (consisting of a number referring to the relevant quadrant and a letter reflecting the relative sizes of the two indices within each quadrant) readily allows one to characterize the data points contained in each triangle: see Table 2.1.[2] For example, the triangle IIIa contains the data points associated with states having negative values of both E and e—thus experiencing negative net income migration while the absolute value of E is larger than that of e. In the same vein, the triangle IVb contains data points associated with states having positive values of E and negative values of e, while the absolute value of E is smaller than that of e—thus states experiencing negative net income migration, etc.

A cursory inspection of Table 2.1 reveals some interesting patterns. First, out of the 51 states, 29 experience positive net income migration, and 22 experience negative net income migration. Second, in 38 out of the 51 states, or roughly 3 in 4, the indices E and e have the same sign so that the net population migration component and the differential income component work in the same direction: in 22 out of the 29 states experiencing positive net income migration and in 16 out of the 22 states experiencing negative net income migration. By contrast, in the remaining 13 states, or roughly 1 in 4, the two indices E and e have opposite signs

[1]It is customary to multiply the raw values of the index E of population migration effectiveness by 100, and thus the same custom is adopted here with regard to the relative per capita income index e.

[2]As could be expected, each of the eight triangles in Fig. 2.1 contains the same number of states as the corresponding triangle in Plane's Fig. 2.3 which, as already said, it supersedes.

Table 2.1 Typology of the US states for decomposing aggregate change in income migration

Net income migration	Index of population migration effectiveness (E) and relative per capita income index (e): Signs	Relative importance	Location in Fig. 2.1	Number of states
Positive	$E > 0$; $e < 0$	$\lvert E\rvert > \lvert e\rvert$	IVa	7
" "	$E > 0$; $e > 0$	" "	Ia	19
" "	" "	$\lvert E\rvert < \lvert e\rvert$	Ib	3
" "	$E < 0$; $e > 0$	" "	IIb	0
Negative	" "	$\lvert E\rvert > \lvert e\rvert$	IIa	4
" "	$E < 0$; $e < 0$	" "	IIIa	12
" "	" "	$\lvert E\rvert < \lvert e\rvert$	IIIb	4
" "	$E > 0$; $e < 0$	" "	IVb	2
				–
				51

so that the net population migration component and the differential income component work in opposite directions: in 7 of the 29 states experiencing positive net income migration and in 6 out of the 22 experiencing negative net income migration.

Aside from comparing the mean average of the absolute value of E (10.3) with that of the absolute value of e (3.5), insights into the impact of the E and e indices on net income migration can be obtained by plotting the E and e indices, separately, against their sum $E + e$, here labeled standardized index of net income migration. In Fig. 2.2, which plots E against $E + e$, the data points tend to congregate around the southwest-northeast diagonal, suggesting that the value of $E + e$ is more or less similar to that of E (a simple regression of E against $E + e$ yields a 1.15 slope coefficient for a R-square of 0.94). By contrast, in Fig. 2.3, which plots e against $E + e$, the data points congregate around the vertical axis, suggesting that the value of $E + e$ is much less influenced by e (a simple regression of e against $E + e$ yields a slope of 2.36 for a R-square of 0.66).

As a result, the role of the relative per capita income index e is to slightly reinforce the impact of the index E of population migration effectiveness in 31 out of the 38 states with indices of identical signs or to counteract slightly the latter in 11 out the of 13 states with indices of opposite signs. In just nine states is the absolute value of the relative per capita income index e larger than that of the index E of population migration effectiveness: seven states where the two indices have an identical sign but only two where they have opposite signs so that net income migration has a sign opposite to that the index E of population migration effectiveness. These two states are Minnesota (where $E = 5.2$ and $e = -5.6$) and Oklahoma (where $E = 2.2$ and $e = -3.5$).

To summarize, the net population migration and differential income components of net income migration are simply reflected by the index E of population migration effectiveness and the relative per capita income index e, respectively, to which they are both proportional. It appears, however, that the former plays a greater role than

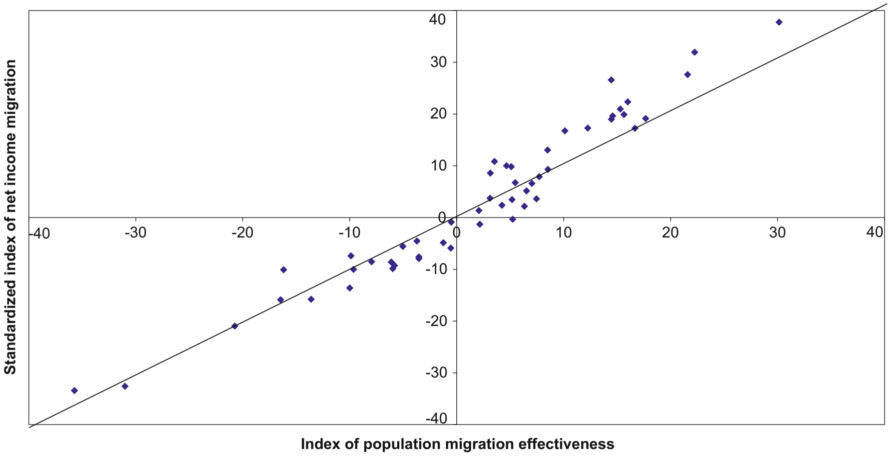


Fig. 2.2 Net income migration versus its net population migration component, US states, 1993–1994

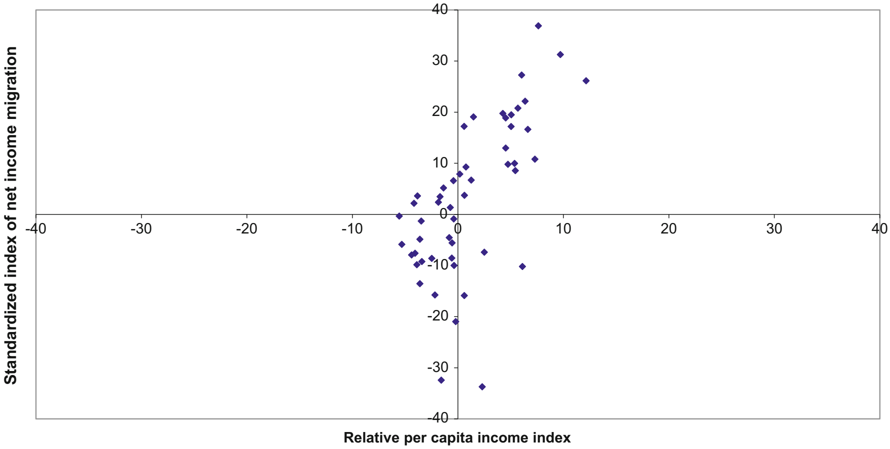


Fig. 2.3 Net income migration versus its differential income component, US states, 1993–1994

the latter in shaping aggregate net income migration. First, in all states but two, the sign of the index E of population migration effectiveness determines the sign of net income migration. Second, in 42 out of the 51 states, its absolute value is generally higher than that of the relative per capita income index.

2.3 The Effectiveness of Income Migration to a Region and Its Constituent Population and Income Elements

The second tool set forth by Plane in his pioneering paper is an index assumed to reflect the effectiveness of aggregate income migration. Borrowing from the definition of the index reflecting the effectiveness of population migration, he defines this index as the ratio of aggregate net income migration to aggregate total income, the latter being equal to the sum of the aggregate in- and outmigrant incomes:

$$F = \frac{Y_I - Y_O}{Y_I + Y_O}.$$

But Plane limits himself to calculating and describing the values taken this index over the US states. It is, however, possible to gain some additional insights into this index. To see this, note that the numerator $Y_I - Y_O$ of F is none other than Y_N so that

$$Y_I - Y_O = \frac{i+o}{2}(I-O) + (i-o)\frac{I+O}{2}.$$

Then on substituting $-O$ for O and similarly $-o$ for o in the latter equation, it readily follows that the denominator $Y_I + Y_O$ of F is

$$Y_I + Y_O = \frac{i+o}{2}(I+O) + (i-o)\frac{I-O}{2}$$

so that

$$F = \frac{\frac{i+o}{2}(I-O) + (i-o)\frac{I+O}{2}}{\frac{i+o}{2}(I+O) + (i-o)\frac{I-O}{2}}.$$

Finally, dividing both numerator and denominator by $\frac{i+o}{2}(I+O)$ yields

$$F = \frac{\frac{I-O}{I+O} + \frac{i-o}{i+o}}{1 + \frac{I-O}{I+O}\,\frac{i-o}{i+o}}$$

or

$$F = \frac{E+e}{1+Ee}.$$

The index F of income migration effectiveness is thus equal to the sum of the index E of population migration effectiveness and the relative per capita income index e divided by 1 plus the product of these two indices. Because the (raw values

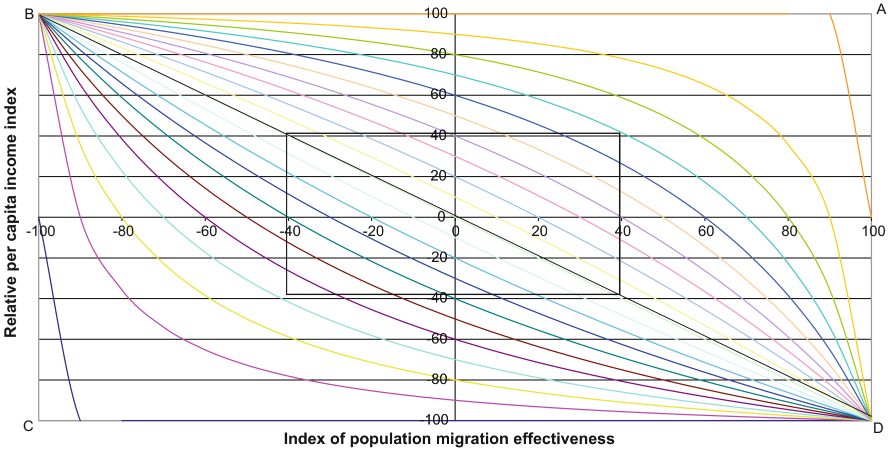


Fig. 2.4 Isocurves of income migration effectiveness, US states, 1993–1994

of the) two indices are comprised between -1 and $+1$, their product is as well so that the denominator of F, $1 + Ee$, is necessarily positive, and as a result F has the same sign as $E + e$.

How does F vary with E and e? One possibility is to trace, with respect to a system of two axes in which the horizontal axis refers to the values of E and the vertical axis to the values of e, a set of isocurves—that is, curves that link points associated with identical values of F. Figure 2.4 shows such a set of isocurves associated with a value of F ranging from -100 to $+100$ in increments of 10.[3]

In practice, the isocurves are simply constructed on the basis of the expression of the e index as a function of the E index, in which the F index is considered as given. To obtain this expression, start from the expression of F rewritten as

$$F(1 + Ee) = E + e$$

or

$$e(1 - FE) = F - E$$

so that

$$e = \frac{F - E}{1 - FE}.$$

[3]Actual values of the index F of income migration effectiveness are also reported using a multiplicative factor of 100.

Figure 2.4 consists of a series of arcs that, starting with the diagonal BD of the ABCD square for $F = 0$, bend more and more upward as F decreases toward -100, when it becomes confounded with the CD side of the ABCD square, but bend more and more downward as F increases toward 100, when it becomes confounded with the BA side of the ABCD square. Would the data points corresponding to the US states be displayed on it, they would be located in the rather contained area inside the square appearing at the center. Indeed, the values of the E index range from -35.7 (California) to 30.1 (Nevada) and those of the e index from -5.6 (Minnesota) to 12.1 (Florida), whereas the values of the F index are comprised between -33.7 (California) and 36.9 (Nevada).

Earlier, the index F of income migration effectiveness was shown to have the same sign as $E + e$. But more remarkably it takes on values that differ from $E + e$ in a predictable manner. Starting with the absolute difference between $(E + e)$ and F

$$(E+e) - F = (E+e) - \frac{E+e}{1+Ee} = (E+e)\frac{Ee}{1+Ee},$$

it readily follows that the relative difference between the two is given by

$$\frac{(E+e) - F}{F} = \frac{(E+e)\frac{Ee}{1+Ee}}{\frac{E+e}{1+Ee}} = Ee.$$

Then the relative error made in substituting $E + e$ for F is simply equal to Ee which varies between -1 and $+1$. Although it could be quite large, in most applied situations, it tends to cover only a narrow interval of its potential range of variation. For example, with respect to the US states, the values of Ee range from a minimum of 99/10,000 or 0.9% to a maximum of 230/10,000 or 2.3%, so that the F index does not differ much in value from the sum of the E and e indices. The relative difference between the two is at most equal to 2.3% in absolute value.

In practice then, F barely differs from E + e, which in Fig. 2.4 is attested in a rather conspicuous fashion: the sections of the isocurves located within the square circumscribing the data points for all of the US states resemble segments of straight lines that are parallel to the (BD) diagonal—that is, with a -1 slope or, equivalently, such that E + e takes on a given value.

To summarize, in most empirical applications, the index F of income migration effectiveness is approximately equal to the sum of index E of population migration effectiveness and the relative per capita income index e.

2.4 The Decomposition of a Region's Change in Per Capita Income into an Inmigrant- and an Outmigrant-Related Component

The third tool introduced by Plane in his pioneering paper is a decomposition technique centered on a region's per capita income, the intent of which is to better understand how in- and outmigration flows affect a region's per capita income.

This technique originates with an equation expressing the change in per capita income. If S is the number of stayers over the observation period and s is the per capita income relating to them, it readily follows that the aggregate income of the end-of-the-period population is equal to $sS + iI$ and the corresponding per capita income $y_E = \frac{sS+iI}{S+I}$. Similarly, the aggregate income of the beginning-of-the-period population is equal to $sS + oO$ and the corresponding per capita income to $y_B = \frac{sS+oO}{S+O}$ so that the per capita income change over the observation period attributable to migration is equal to

$$\Delta y = y_E - y_B = \frac{sS + iI}{S + I} - \frac{sS + oO}{S + O}.$$

From there, Plane decomposes Δy into three terms (Plane 1999: 207), which is more or less suitable for interpretation. It is, however, possible to set forth an alternative decomposition into two terms specifically tied to the in- and outmigrant flows, which happens to simplify as well as clarify the interpretation. To see this, substitute $(S + I) - I$ for S in the numerator of y_E and similarly $(S + O) - O$ for S in the numerator of y_B so that

$$\Delta y = \frac{s[(S + I) - I] + iI}{S + I} - \frac{s[(S + O) - O] + oO}{S + O}.$$

This yields

$$\begin{aligned}\Delta y &= \frac{s(S + I) + (i - s)I}{S + I} - \frac{s(S + O) + (o - s)O}{S + O} \\ &= s + (i - s)\frac{I}{S + I} - s - (o - s)\frac{O}{O + S}\end{aligned}$$

and eventually

$$\Delta y = (i - s)\frac{I}{S + I} - (o - s)\frac{O}{S + O}.$$

In other words, the change in per capita income attributable to migration is simply the difference between an immigrant-related component, equal to the product of the per inmigrant/stayer capita income differential by the proportion of immigrants in

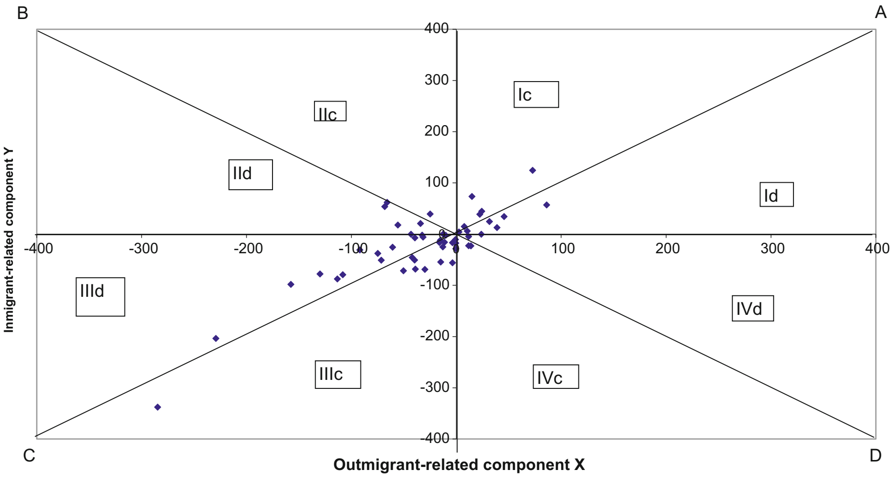


Fig. 2.5 Decomposition of per capita income change into its out- and inmigrant-related components, US states, 1993–1994

the end-of-the-period population, and an outmigrant-related component, equal to the product of the outmigrant/stayer per capita income differential by the proportion of outmigrants in the beginning-of-the-period population.

Based on this, Fig. 2.5 reveals the impact of the two differential income variables on per capita income change for the US states. It was obtained by a slight modification of the axes in Plane's Fig. 2.4. Instead of referring to the values of the two per capita income differences between in- and outmigrants on the one hand and stayers on the other, the axes refer to the in- and outmigrant-related components, respectively. Thus the horizontal axis refers to the values of the outmigrant-related component:

$$X = (o - s)\frac{O}{S + O},$$

while the vertical axis refers to the values of the inmigrant-related component:

$$Y = (i - s)\frac{I}{S + I}.$$

Naturally, all of the states experiencing a gain in per capita income attributable to migration are such that $Y - X > 0$, and thus they are located above the southwest to northeast diagonal (or CA) of the ABCD square.

Unlike Plane's Fig. 2.4, Fig. 2.5 above enables one to display the 51 data points without resorting to any approximation that stems from the presence of a third distracting component such as in Plane's original framework, but the two figures

Table 2.2 Typology of the US states for decomposing per capita income change attributable to migration

Per capita income change	In- and outmigrant components (X and Y, respectively) Signs	Relative importance	Location in Fig. 2.5	Number of states
Positive	$X > 0$; $Y > 0$	$\|X\| < \|Y\|$	Ic	5
" "	$X < 0$; $Y > 0$	" "	IIc	1
" "	" "	$\|X\|>\|Y\|$	IId	5
" "	$X < 0$; $Y < 0$	" "	IIId	14
Negative	" "	$\|X\| < \|Y\|$	IIIc	16
" "	$X > 0$; $Y < 0$	" "	IVc	2
" "	" "	$\|X\| > \|Y\|$	IVd	2
" "	$X > 0$; $Y > 0$	" "	Id	6
				–
				51

cannot be easily compared.[4] Figure 2.5 is, however, rather similar to Plane's Fig. 2.2 because the coordinates of any point in the former figure are identical to those of the corresponding point in the latter, up to a positive multiplicative factor.[5]

Just as in Sect. 2.2, let us label the quadrants defined by the two axes I to IV when moving counterclockwise from the quadrant with positive values of both X and Y (quadrant I) toward the quadrant with positive values of X and negative values of Y (quadrant IV). Moreover, in each of the quadrants I–IV, let us join the origin to the pertinent corner on the ABCD square (A, B, C, and D, respectively), and, among the eight ensuing triangles, let us assign the letter c to those in which the absolute value of X is smaller than the absolute value of Y and the letter d to those in which it is larger.

Again, the label assigned to each triangle, consisting of a number referring to the relevant quadrant and a letter reflecting the relative absolute sizes of X and Y, readily allows one to characterize the data points that it contains: see Table 2.2 from which emerge some interesting patterns.

First, out of the 51 states, 25 experience a positive change in per capita income, and 26 experience a negative one. Second, in 41 out of the 51 states, or roughly 4 in 5 states, the X and Y variables have the same sign so that they work in opposite directions (recall that change in per capita income is given by $Y - X$): 19 out of the 25 states experiencing a positive change in per capita income and 22 out of the 26 states experiencing negative change in per capita income. By contrast, in the

[4] All the more so because the axes have been reversed and that the outmigrant/stayer per capita income differential is obtained by subtracting the per capita income of stayers from that of outmigrants, not the other way around as in Plane (1999).

[5] The number of data points is necessarily identical in corresponding quadrants, but their distribution among the two triangles that constitute each quadrant may differ as a result of the relative values of the in- and outmigrant proportions in the end- and beginning-of-the-period populations.

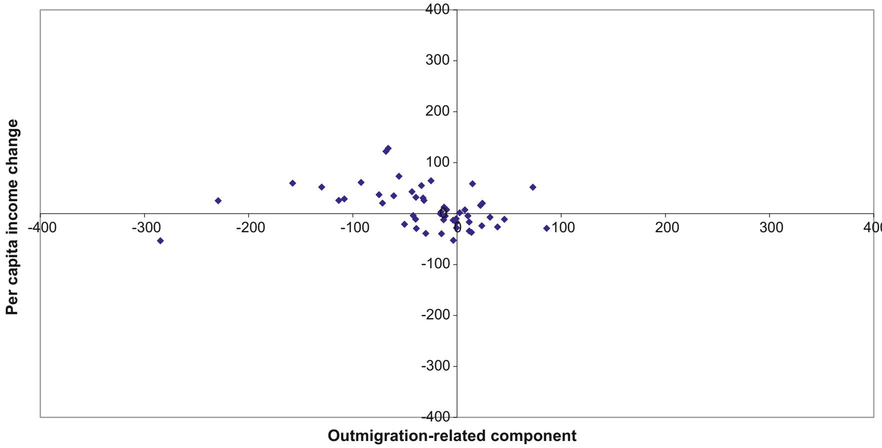


Fig. 2.6 Per capita income change versus it outmigrant-related component, US states, 1993–1994

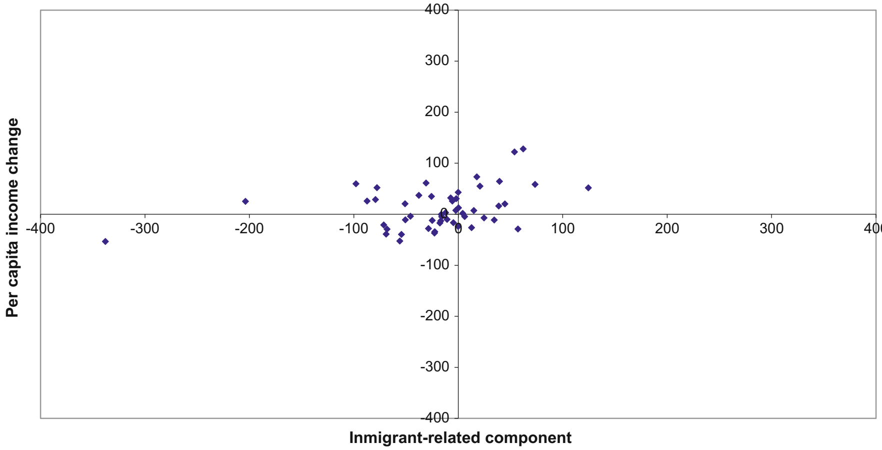


Fig. 2.7 Per capita income change versus its inmigrant-related component, US states, 1993–1994

remaining 10 states, or roughly one in 5 states, X and Y have opposite signs so that they work in the same direction: in 6 of the 25 states experiencing positive change in per capita income and 4 of the 26 experiencing negative change in per capita income.

Aside from comparing the mean averages of their absolute values (47.8 and 44.8, respectively), additional insights into the impact of the X and Y variables on change in per capita income can be obtained by plotting them separately against their difference. At first glance, the data points in the two figures are scattered around the origin. In Fig. 2.6 which plots X against $Y - X$, they tend to be oriented along a northwest-southeast direction as confirmed by a simple regression of X against $Y - X$ (slope coefficient equal to –0.14 for a R-square equal to 0.06). By contrast, in Fig. 2.7 they tend to be oriented along a southwest-northeast direction as confirmed by a simple regression of Y against $Y - X$ (slope coefficient equal to 0.20 for a R-square equal to 0.24).

To summarize, the Y and X variables—that is, the two variables linked to the differences in per capita income exhibited by in- and outmigrants on the one hand and the stayers on the other—appear to play roughly a similar role in shaping change in per capita income.

2.5 Recapitulation

This chapter has attempted to unravel the tools introduced by Plane (1999) for examining the impact of interregional migration flows on the aggregate amount of money thus redistributed. In doing so, it has succeeded in rendering more meaningful Plane's decomposition techniques pertaining to a region's changes in aggregate as well as per capita income as a result of people moving between regions:

- Change in a region's aggregate income can be decomposed into a net population migration component and a differential income component which are in tune with the index E of population migration effectiveness and the relative per capita income index e, respectively, of which, it turns out, the former component is highly dominant over the latter one.
- Change in a region's per capita income can be expressed as a difference between an immigrant-related variable (equal to the inmigrant/stayer per capita income differential times the inmigrant proportion in the end-of-the-period population) and an outmigrant-related variable (equal to the outmigrant/stayer per capita income differential times the outmigrant proportion in beginning-of-the period population), of which neither one appears to have a dominant role.

Finally, Plane's income equivalent to the measure of effectiveness of population migration to a region widely used in migration research may not be as promising a notion, because, in practice, his measure of effectiveness of income migration to a region turns out to take on values that do not differ very much from the sum of the index of population migration effectiveness and the index of relative per capita incomes among in- and outmigrants.

Reference

Plane D (1999) Geographical pattern analysis of income migration in the United States. Int J Popul Geogr 5:195–212

Chapter 3
A Short Exercise to Assess the Effects of Temporal and Spatial Aggregation on the Amounts of Spatial Spillovers

Sungyup Chung and Geoffrey J. D. Hewings

Abstract In the modifiable areal unit problem (MAUP) literature, it is known that in measuring spillover effects, the magnitudes of spillovers are different depending on the scale of the temporal/spatial units. However, the research has not addressed whether the magnitude of the spillover increases or decreases as the unit of measurement increases. From an exercise using a constructed regional economic system, it is shown that depending on how researchers make assumptions about the data generating process of regional economies, the magnitude of the measured spillover may be smaller or larger as the unit of measurement increases. It is also argued that with a reasonable assumption that there is a common factor affecting the data generating processes of regional units (i.e., the regional economic system is characterized by a multi-level structure), the larger the unit of scale, the smaller the amount of spillover.

Keywords Temporal aggregation · Spatial aggregation · Multi-level structure · Spillover effects

JEL Classifications R10 · R12 · C43 · C63

The research was performed while resident in the Regional Economics Applications Laboratory and Department of Economics, University of Illinois, Urbana, IL, 61801-3671

S. Chung
Industry and Labor Research Team, Economic Research Department, Seoul, South Korea

G. J. D. Hewings (✉)
Regional Economics Applications Laboratory and Department of Economics, University of Illinois, Urbana, IL, USA
e-mail: hewings@illinois.edu

R. S. Franklin (ed.), *Population, Place, and Spatial Interaction*, New Frontiers in Regional Science: Asian Perspectives 40,
https://doi.org/10.1007/978-981-13-9231-3_3

3.1 Introduction

This chapter aims to unravel how the magnitude of spatial spillovers changes as the frequency/scale of the temporal/spatial units is altered. The unit of scale, whether temporal or spatial, has been a problem in the empirical analysis of spatial economics.[1] For example, the implications are usually different when analysis is conducted using US state level data or when using US county level data. The issue becomes even more important when measuring the spillover effects from neighboring regions; the magnitude of the spillovers changes with the spatial unit of the observations.

In this chapter, the size of the spillover effects is defined using forecast error variance decomposition (FEVD). FEVD decomposes the variances of the regional behavior into their sources for a given forecasting horizon. For example, the portion of the neighborhood innovation relative to the regional variance can be quantitatively calculated using FEVD.[2]

Given a multi-level structural representation of an economic system,[3] one can assume that there is a common factor (e.g., at the national level) governing the behavior of the regional economic system, we should expect to observe a smaller volume of spillovers with larger frequency (scale) of temporal (spatial) units. Although these conclusions are derived from an experiment using a simple geographically defined regional economic system, this finding is more in concord with our expectations on real-world data. Under the assumption that there is a common factor in the regional economic system, we should expect less spillover effects with larger scale of spatial or temporal units. Over a century ago, W. F. Gosset (student) also thought about this problem, and in his letter to Karl Pearson in December 1910, he made a similar conclusion: “Now in general, the correlation weakens as the unit of time or space grows larger and I can’t help thinking that it would be a great thing to work out the law according to which the correlation is likely to weaken with increase of unit” (Pearson 1990). His notion of correlation can be translated into spillovers in the regional economic activity analysis context.

The assumption that there is a common factor governing the behaviors of the regional economic system plays an important role in assessing the neighborhood effects. If it is assumed that there is no such thing as a common factor, and regional

---

[1] Modifiable areal unit problem (MAUP) is well addressed in Clark and Avery (1976), Fotheringham and Wong (1991), Haitovsky (1973), Openshaw and Taylor (1979), Prais and Aitchison (1954), and Wong (2004).

[2] Also, since cumulative impulse response function (CIRF) measures the cumulative effect of a regional shock on the future values of regional values, the “direction” of a spillover can be defined as the sign of the long-term cumulative response. An exercise on the CIRFs is also considered here, but it is found that CIRF results do not exhibit any pattern that can be generalized on a geographically defined economic system. The results are presented in Appendix 1. All appendices are available at www.real.illinois.edu/d-paper/15/Aggregation%20Appendix.pdf

[3] Contrary to the multi-level structure regional economic system, a single-level structure regional economic system does not consider the existence of the higher-level common factor. More details can be found in Chung and Hewings (2015).

economic behaviors are governed only by their own forces and those of their neighbors, the implication is opposite. In other words, as the unit of spatial scale becomes larger, we should expect to observe a larger volume of spillovers with larger frequency/scale of temporal/spatial units. Depending on how researchers make assumptions about the data generating process of regional economies, the magnitude of the measured spillover may be smaller or larger as the unit of measurement increases.

In many cases in empirical analysis, it is somewhat reasonable to assume the existence of a common factor in a regional economic system, since it would be difficult to imagine a regional economic system that is closed to outside influences. For example, the states in the USA are affected by the US national level economic shocks, and the counties of a state are affected by the states' economic shocks such as state government policies. The concept of common factor (a multi-level concept) is well addressed in Bai and Wang (2012) and Corrado and Fingleton (2011). In practice, the region common factor is the national level economic behavior, and in our exercise, it is an exogenous, or almost exogenous, factor affecting the regional economic behavior.[4] By introducing a hierarchical structure into spatial analysis, a multi-level analysis argues that the co-moving behaviors of spatial units are largely due to these higher-level determinant(s). On the contrary, under the assumption that there is no such thing as common factor, a single-level analysis should be conducted, and co-moving behaviors are largely due to the spillover effects. By comparing the forecast error variance decomposition and the impulse responses of regional observations using both concepts, this chapter argues that a multi-level concept explains the real-world data in a more realistic way.

Recent studies and a small exercise on the spatial or temporal aggregation problems are briefly introduced in the next section. In Sect. 3.3, an artificial regional economic system is constructed to see how the observed spillovers change with the level of temporal/spatial aggregation, and Sect. 3.4 provides some concluding remarks.

3.2 Spillovers and Temporal/Spatial Aggregation in Practice

While most of the previous work has focused on other issues such as the biasedness or the efficiency loss of the estimated coefficients at the aggregated level when the true data generating process is defined at the micro level, the primary focus of

[4] Almost exogenous means regional shock does not have a significant impact on the region common economic behaviors. More practically, in the model structure, the coefficients associated with the effect from regional shock to the region common behavior should be close to zero so that the local impact on the global behavior decays very fast throughout time. Conceptually, a multi-level structure model with endogenous region common factor can be regarded as a single-level structure model in our exercise.

this chapter is on what we should expect to observe in terms of the spillover effects with different levels of data aggregation.[5]

However, attempts to find a general “law” or some monotonic relationships between the scale of grouping or the level of aggregation and the spillover effects, which can be defined in terms of forecasting error variance decomposition and the impulse response function, are very rare. There is some empirical work exploring the scale of unit effects on the correlation coefficients or Moran’s *I* (see Arbia 1989). Gehlke and Biehl (1934) explored the effect of the aggregation on the correlation coefficient between observations, and Smith (1938) explored the correlation coefficient with different plot sizes in an agricultural experiment. Also, Dusek (2004) showed how different geographical statistics varied with different levels of aggregations on Hungarian regional economic data. However, Gehlke and Biehl’s (1934) finding that the correlation coefficient for variables of absolute measurement increases when areal units are aggregated contiguously and Dusek’s (2004) finding that more aggregated observations exhibit higher Moran’s *I* are exactly the opposite of Gosset’s idea that the correlation will weaken with the increase of scale.

In empirical studies in agricultural economics, Gosset’s idea is supported in both theoretical and empirical work. For example, in Gelfand et al. (2010), when the covariance between spatial units are defined in terms of area and distance, the inverse relationship between the aggregation level and the covariance is easy to assess. Following Gelfand et al. (2010), consider a stationary spatially continuous stochastic process, $S(x)$, with covariance function $\mathrm{cov}\{S(x), S(y)\} = \sigma^2\rho(u)$ where u is the distance between the locations x and y.

The covariance between spatial averages of $S(\cdot)$ over two regions A and B is

$$\gamma(\mathrm{A},\mathrm{B}) = (|\mathrm{A}| \times |\mathrm{B}|)^{-1}\sigma^2\int_{\mathrm{A}}\int_{\mathrm{B}}\rho(\|x-y\|)\,dx\,dy \tag{3.1}$$

where $|\cdot|$ denotes area and $\|\cdot\|$ distance.

Equation (3.1) supports Gosset’s statement since in a space where the correlations are decreasing with the distance, the correlation weakens with the larger scale of unit. For instance, when the covariance is proportional to the inverse distance, then the aggregation of the four equal-sized spatial units will result in the half size value of the

[5]Brewer (1973), Wei (1981), and Weiss (1984) tackle the issues related to temporal aggregation in empirical studies. Marcellino (1999) reviewing the literature on this issue showed that impulse response functions and forecast error variance decompositions, along with other properties such as Granger-causality and cointegration, change with the level of aggregation. On the other hand, the spatial aggregation problem, or the modifiable areal unit problem (MAUP), has also been an interest of spatial analysts for a long time since the pioneering work of Gehlke and Biehl (1934). The effects on standard regression estimators are addressed in Barker and Pesaran (1990), Okabe and Tagashira (1996), Tagashira and Okabe (2002), and Griffith et al. (2003). Their main findings are that “the GLS estimators of regression’s parameters are BLUE with a sampling variance greater than that obtained using GLS on the original data” (Arbia and Petrarca 2011). Arbia and Petrarca (2011) also explored the efficiency loss of the estimators in the presence of spatial dependency.

original covariance. Empirical study of the correlations defined by the distance between objects can be found in Whittle (1956, 1962) and McCullagh and Clifford (2006).

In more general spatial economic analysis, where the spillovers are not defined by correlations, and negative spillover exists, spatial dependency should be assessed using the concept of spillover that can be defined with, for example, forecast error variance decomposition or impulse response function, rather than correlation coefficients. In other words, the correlation coefficient measures the degree of the co-movement, whereas the spillover effects measures the effects from the neighborhood.[6]

This conceptual distinction is more straightforward when we compare one regional economic system that has a common factor governing the region's subspatial units (a multi-level structure regional economic system) and the other regional economic system without any common factor (a single-level structure regional economic system). For example, if there is a common factor, two independent counties can have a high correlation coefficient if they are exposed to a same state-level shock and have similar response functions in response to this shock. Thus, if an economy is a multi-level structure, regions with high correlation coefficients may not be highly connected. That is, they are correlated to a common factor rather than directly dependent on each other. On the contrary, of course, if there is no common factor, a high correlation coefficient implies large neighborhood spillover effects.

In practice, we cannot distinguish whether the observed high correlations between regional units are due to the existence of a common factor or the high dependency between regional units. Thus, the researcher's belief in the existence of a common factor becomes more important in empirical analysis. For example, equating the co-movement to spillovers where the economy is a multi-level structure can exaggerate the impact of local policy on its neighborhood regions. Likewise, assuming the existence of a common factor where there is no such thing can underestimate the impact of local neighborhood shock. In this chapter, the constructed regional economic system is assumed to be a multi-level structure considering the higher-level shocks that are common to the regional economies, such as monetary/fiscal policy or international commodity price shocks. A more detailed discussion about region common factor can be found in Chung and Hewings (2015).

[6]Since the spillover effects varies not only with space but also with time, the study of spillover effects should be conducted in terms both spatial and temporal scale.

3.2.1 Experimental Analysis to Explore How the Amount of Spillover Varies with the Measurement Units

As a motivation to assess the consequences of the spatiotemporal aggregation of spatial units in a multi-level regional economic system, an exercise was conducted using the approach of Bai and Wang's (2012) multi-level dynamic factor analysis on selected regions. The exercise shows that on average, the portion of spillover effects decreases with larger spatial scale or with higher frequency of temporal units. Of course, it should be admitted that the comparison of the results from the models with different scales is problematic since the underlying assumptions for the data generating process of each set of observations may be different. Nevertheless, the results provide some vague pictures about the amount of spillovers with different scales of observation units. Also, the framework provided by this model will be used again in our main exercise on the constructed regional economic system.

An example of Bai and Wang's (2012) multi-level dynamic factor model is shown in Eqs. (3.2), 3.3, and (3.4):

$$\begin{bmatrix} y_t^1 \\ \vdots \\ y_t^R \end{bmatrix} = \begin{bmatrix} \delta^1(L) & \gamma^1(L) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \delta^R(L) & 0 & \dots & \gamma^R(L) \end{bmatrix} \begin{bmatrix} g_t \\ f_t^1 \\ \vdots \\ f_t^R \end{bmatrix} + \begin{bmatrix} u_t^1 \\ \vdots \\ u_t^R \end{bmatrix} \tag{3.2}$$

$$D^r(L)u_t^r = \varepsilon_t^r \quad \forall r \in \{1, \dots R\} \tag{3.3}$$

$$\begin{bmatrix} g_t \\ f_t^1 \\ \vdots \\ f_t^R \end{bmatrix} = \begin{bmatrix} \varphi_{11} & \cdots & \varphi_{1\,(R+1)} \\ \vdots & \ddots & \vdots \\ \varphi_{(R+1)1} & \cdots & \varphi_{(R+1)(R+1)} \end{bmatrix} \begin{bmatrix} g_{t-1} \\ f_{t-1}^1 \\ \vdots \\ f_{t-1}^R \end{bmatrix} + \begin{bmatrix} \eta_t^g \\ \eta_t^1 \\ \vdots \\ \eta_t^R \end{bmatrix} \tag{3.4}$$

where

y_t^r is $p \times 1$ vector of endogenous regional observations,
f_t^r is unobservable fundamental forces that affect the dynamics of y_t^r, and
g_t is unobservable fundamental force that affects the dynamics of $\left(y_t^1, \dots, y_t^R\right)$

$E\left(\eta_t^r \mid g_{t-1}, f_{t-1}^1, \dots, f_{t-1}^R\right) = 0 \ \forall t \in \{1, \dots, T\}$ and $\forall r \in \{g, 1, \dots, R\}$,
$E\left(u_t^r \mid g_t, f_t^1, \dots, f_t^R\right) = 0 \ \forall t \in \{1, \dots, T\}$ and $\forall r \in \{1, \dots, R\}$,
$E\left(u_t^{ri} u_t^{rj} \mid g_t, f_t^1, \dots, f_t^R\right) = 0 \ \forall i \neq j \in \{1, \dots, p\}, \forall t \in \{1, \dots, T\}$ and $\forall r \in \{1, \dots, R\}$, and
$E\left(u_t^a u_t^b \mid f_t^1, \dots, f_t^R\right) = 0 \ \forall t \in \{1, \dots, T\}$ and $\forall a \neq b, r \in \{1, \dots, R\}$.

The observed variables in each region, y_t^r, are decomposed into three parts of shocks: (1) region common factor, g_t; (2) regional dynamic factor, f_t^r; and (3) idiosyncratic error, u_t^r. The spatiotemporal relationship of the factors can be found in the coefficient matrix of the state equation of the dynamic factor (Eq. 3.4). From this coefficient matrix, we can derive the forecast error variance decomposition (FEVD) and assess the spatiotemporal dynamics of regional business cycles. Thus, in this sense, the "spillover" is larger if the neighborhood region's percentage portion in FEVD is larger.

Using Gibbs sampling algorithm in WinBUGS,[7,8] FEVD are estimated on the selected regional employment series with different temporal frequencies and spatial scales: monthly frequency county level, group level, state level, and regional division level data are analyzed, and the county level monthly, quarterly, and biannual frequency data are also analyzed. At the county level, Peoria, Tazewell, McLean, Champaign, and Vermilion Counties in Illinois were selected, since outside this group of counties, population is very sparse, so those five counties are thought to have a natural common border. At the group level, those five counties are referred to as the "I74" group.[9] Other groups are the St. Louis Group, Quad City Group, Springfield Group, and Chicago Group that are all located inside or close to the state of Illinois. Each group consists of 2–12 counties, and, like "I74" group, populations are very sparse outside the counties constituting each group. The detailed descriptions are provided in Fig. 3.1.

At the state level, six Great Lake states, Illinois, Indiana, Michigan, Minnesota, Ohio, and Wisconsin, were selected. The regional division level units are simply the four US regional divisions (Northeast, Midwest, South, and West).[10] To compare the consequences of the use of different temporal frequencies of observations, county level monthly data were aggregated to generate quarterly and biannual series.

Using those data sets, multi-level dynamic factor models (Eqs. 3.2, 3.3, and 3.4) were estimated. However, it should be admitted that it is not appropriate to compare the results of those exercises since each model with different scales assumes a different definition of the region common factor. Thus, the results are only useful in terms of providing a rough idea of how the magnitude of spillovers change with different observation scales. The FEVD results with different temporal scales derived from a multi-level model are presented in Table 3.1.[11]

---

[7]Bayesian inference using Gibbs sampling for Windows.

[8]Only VAR(1) structure model estimation results are presented in Sect. 3.2 since, according to the deviations information criteria (DIC), any higher lag order does not outperform VAR(1) structure of equations introduced in Sect. 3.1. Also, the Bayesian inference relies on priors, but for this exercise, uninformative priors were used. More detailed procedures can be found in Chung and Hewings (2015).

[9]It is named after the interstate highway I74, because those counties are located along this highway.

[10]Except for the regional division level spatial units, the selection of the regional units at all other levels suffer borderline problem since there is a possibility that some relevant regional unit could have been omitted. However, further consideration of solving this borderline problem was not tried here since this section aims to sketch how the aggregation of spatial temporal unit affects the observed spillover effects in the real-world data and does not aim to exactly identify the regional economic system.

[11]The results for single-level model are also available in Appendix 2.

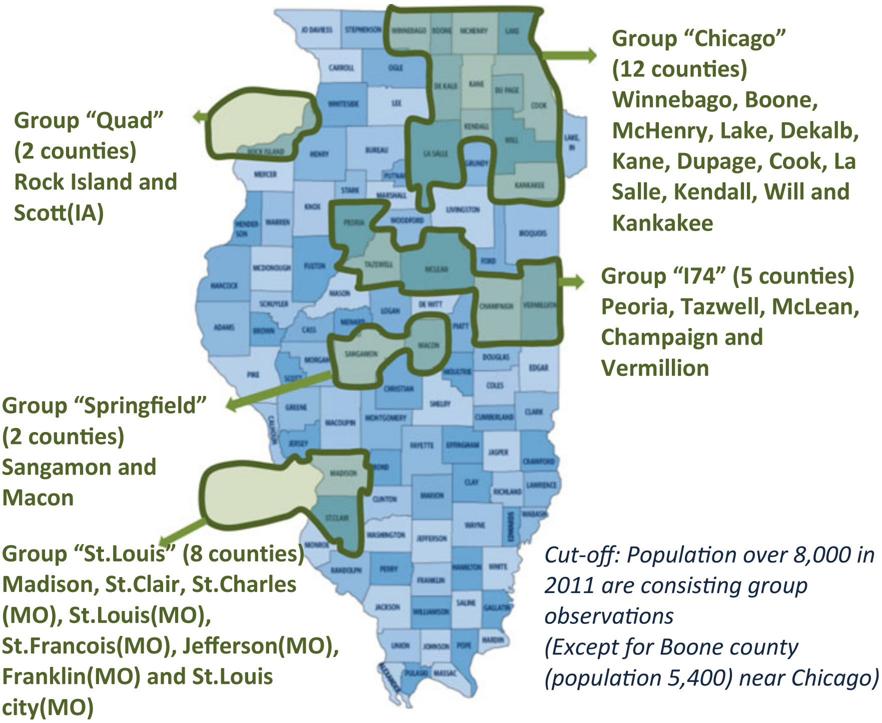


Fig. 3.1 Group data specifications

The point estimates of 1-year ahead FEVD show that although the amount of spillover varies a great deal depending on the individual regional units, on average, larger temporal scale observations generate a smaller neighborhood portion of the FEVD. Accordingly, as the frequency of the temporal scale becomes smaller, the portion from the innovation from the own region decreases, but the innovation from the region common factor increases.

The FEVD results with different spatial scales exhibit a similar trend. In Table 3.2, the portion of the neighborhood innovation decreases with the larger scale of spatial units on average, and the portion of the region common factor increases.

The regional economic performance attributable to a neighborhood or a region common factor is thought to be dependent on its geographical location or the relative position in the value chain, but the portion on average seems to exhibit some kind of monotonic trends related to the level of temporal/spatial aggregation in a multi-level structure model. If there really exist trends, then they should concur with the thought

Table 3.1 Temporal aggregation using a multi-level structure model: point estimates of 1-year ahead forecast error variance decomposition

County level—monthly frequency										
	A	B	C	D	E	G	Total	Neighbor	Own	Common
A	23.4%	65.7%	2.8%	3.7%	3.1%	1.4%	100%	75.2%	23.4%	1.4%
B	18.3%	69.8%	4.1%	3.6%	2.6%	1.6%	100%	28.6%	69.8%	1.6%
C	15.2%	75.9%	1.2%	3.9%	3.3%	0.5%	100%	98.2%	1.2%	0.5%
D	32.3%	43.2%	16.3%	3.0%	1.0%	4.2%	100%	92.8%	3.0%	4.2%
E	28.4%	41.1%	13.2%	3.1%	3.9%	10.3%	100%	85.8%	3.9%	10.3%
Average								**76.1%**	20.3%	3.6%
County level—quarterly frequency										
	A	B	C	D	E	G	Total	Neighbor	Own	Common
A	4.3%	1.8%	10.0%	4.7%	1.9%	77.3%	100%	18.4%	4.3%	77.3%
B	2.4%	1.4%	5.8%	3.1%	15.9%	71.4%	100%	27.2%	1.4%	71.4%
C	0.9%	0.4%	21.6%	1.7%	58.6%	16.8%	100%	61.6%	21.6%	16.8%
D	1.8%	0.7%	22.5%	5.9%	55.3%	13.8%	100%	80.2%	5.9%	13.8%
E	1.5%	0.2%	34.6%	2.1%	60.6%	1.0%	100%	38.4%	60.6%	1.0%
Average								**45.1%**	18.8%	36.1%
County level—biannual frequency										
	A	B	C	D	E	G	Total	Neighbor	Own	Common
A	0.2%	0.4%	1.0%	0.0%	0.0%	98.3%	100%	1.5%	0.2%	98.3%
B	0.7%	69.6%	4.4%	0.3%	2.4%	22.5%	100%	7.9%	69.6%	22.5%
C	0.0%	26.9%	1.7%	3.8%	0.0%	67.6%	100%	30.7%	1.7%	67.6%
D	0.6%	84.1%	3.1%	9.5%	2.3%	0.4%	100%	90.1%	9.5%	0.4%
E	2.9%	70.1%	4.5%	6.7%	8.2%	7.7%	100%	84.1%	8.2%	7.7%
Average								**42.9%**	17.8%	39.3%

Capital letters denote regions, for example, "G" denotes region common shock and A, Peoria County; B, Tazewell County; C, McLean County; D, Champaign County; E, Vermilion County

experiment of a constructed regional economic system presented in the next section; further, the real-world results from the multi-level structure model are consistent in most aspects with the results drawn from the exercise on the constructed regional economic system.

In the next section, by constructing an artificial regional economic system, the consequences of the temporal/spatial aggregation over the amount of spillovers are compared with the results from this section's empirical findings.

Table 3.2 Spatial aggregation using a multi-level structure model: point estimates of 12-step ahead forecast error variance decomposition

County level[a]

	A	B	C	D	E	G	Total	Neighbor	Own	Common
A	23.4%	65.7%	2.8%	3.7%	3.1%	1.4%	100%	75.2%	23.4%	1.4%
B	18.3%	69.8%	4.1%	3.6%	2.6%	1.6%	100%	28.6%	69.8%	1.6%
C	15.2%	75.9%	1.2%	3.9%	3.3%	0.5%	100%	98.2%	1.2%	0.5%
D	32.3%	43.2%	16.3%	3.0%	1.0%	4.2%	100%	92.8%	3.0%	4.2%
E	28.4%	41.1%	13.2%	3.1%	3.9%	10.3%	100%	85.8%	3.9%	10.3%
Average								**76.1%**	20.3%	3.6%

Group level[b]

	A	B	C	D	E	G	Total	Neighbor	Own	Common
A	3.1%	7.2%	65.0%	18.6%	0.5%	5.7%	100%	91.2%	3.1%	5.7%
B	1.8%	12.1%	64.0%	18.3%	0.5%	3.2%	100%	84.7%	12.1%	3.2%
C	1.9%	1.2%	84.9%	8.0%	0.2%	3.8%	100%	11.3%	84.9%	3.8%
D	1.4%	6.7%	62.2%	24.4%	1.0%	4.4%	100%	71.3%	24.4%	4.4%
E	2.0%	7.3%	59.5%	21.6%	1.4%	8.1%	100%	90.5%	1.4%	8.1%
Average								**69.8%**	25.2%	5.0%

State level[c]

	A	B	C	D	E	F	G	Total	Neighbor	Own	Common
A	79.1%	2.3%	7.9%	6.2%	0.8%	0.2%	3.5%	100%	17.4%	79.1%	3.5%
B	57.5%	2.4%	29.2%	4.4%	0.4%	0.6%	5.5%	100%	92.1%	2.4%	5.5%
C	7.6%	0.2%	70.3%	10.3%	4.7%	3.1%	3.8%	100%	25.9%	70.3%	3.8%
D	4.8%	0.2%	6.4%	77.5%	3.3%	2.7%	5.2%	100%	17.3%	77.5%	5.2%
E	24.2%	1.4%	3.3%	2.3%	65.4%	0.1%	3.3%	100%	31.3%	65.4%	3.3%
F	8.4%	0.4%	12.4%	46.1%	11.3%	3.7%	17.6%	100%	78.7%	3.7%	17.6%
Average									**43.8%**	49.7%	6.5%

Regional division level[d]

	A	B	C	D	G	Total	Neighbor	Own	Common
A	16.2%	0.2%	4.4%	0.3%	78.9%	100%	4.9%	16.2%	78.9%
B	0.7%	2.1%	0.2%	0.1%	96.9%	100%	1.0%	2.1%	96.9%
C	3.1%	0.3%	2.7%	0.8%	93.1%	100%	4.2%	2.7%	93.1%
D	8.0%	0.2%	2.5%	0.4%	88.9%	100%	10.7%	0.4%	88.9%
Average							**5.2%**	5.3%	89.5%

Capital letters denote regions, for example, "G" denotes region common shock

[a]A, Peoria County; B, Tazewell County; C, McLean County; D, Champaign County; E, Vermilion County

[b]A, St. Louis Group; B, Quad Cities Group; C, Springfield Group; D, Chicago Group; E, I74 Group

[c]A, Illinois; B, Indiana; C, Michigan; D, Minnesota; E, Ohio; F, Wisconsin

[d]A, Northeast Regional Division; B, Midwest Regional Division; C, South Regional Division; D, West Regional Division

3.3 The Constructed Regional Economic System

3.3.1 Assumptions on the Constructed Regional Economic System

The aggregation scheme that we are using on our practice is average sampling,[12] i.e., Aggregated Observation at time $t = \sum$ Disaggregated Observations during time t. Also, the true data generating process is defined at the most disaggregated level.

The first assumption that the aggregation scheme should be average sampling is not applicable to most empirical analyses in its original form, but in many cases, aggregations are approximately average sampling. In the previous section, the dependent variable at the aggregated level is constructed from the aggregation of the disaggregated level data, log-transformed, and first-differenced. Thus, the previous example is not an exact average sampling. However, since the first-differenced value of log-transformed value is a first-order Taylor expansion of the growth rate, and assuming that the initial status of the observations are approximately the same, the arithmetic average of the growth rate is almost the same as the first-differenced log-transformed observations.[13]

The second assumption is that the true data generating process is defined at the most disaggregated level. This assumption is necessary in order to conduct the experiments using a constructed regional economic system since if the true data generating process is defined at an aggregated level, it is not possible to discuss what is going on at the disaggregated level. Additionally, in a real-world situation, spatial interactions are more apparent at the disaggregated level.[14]

In addition to those two restrictions, only a spatially stationary data generating process[15] is employed in this exercise. With a spatially non-stationary process, any local shock will explode the whole regional economic system.

The artificial regional economic system consists of 1024 cities (numbered from 1 to 1024) arranged in a 32 × 32 rectangular country. Four cities comprise 1 county; thus there are 256 counties arranged in a 16 × 16 panel. Likewise, 4 counties comprise 1 group (64 groups), 4 groups comprise 1 state (16 states), and 4 states comprise 1 division (4 divisions). The graphical representation of the land structure is shown in Fig. 3.2.

---

[12] If a time series is stock data, the aggregation of the time series can be a point-in-time sampling, for example, Aggregated Observation at t = Last Disaggregated Observation during t. For more detail, see Marcellino (1999).

[13] For example,

$$\begin{aligned}\ln(a_t + b_t + c_t) - \ln(a_{t-1} + b_{t-1} + c_{t-1}) &\approx \tfrac{a_t - a_{t-1} + b_t - b_{t-1} + c_t - c_{t-1}}{a_{t-1} + b_{t-1} + c_{t-1}} \\ &\approx \tfrac{1}{3}\left(\tfrac{a_t - a_{t-1}}{a_{t-1}} + \tfrac{b_t - b_{t-1}}{a_{t-1}} + \tfrac{c_t - c_{t-1}}{a_{t-1}}\right) \approx \tfrac{1}{3}\left(\tfrac{a_t - a_{t-1}}{a_{t-1}} + \tfrac{b_t - b_{t-1}}{b_{t-1}} + \tfrac{c_t - c_{t-1}}{c_{t-1}}\right) \\ &\approx \tfrac{1}{3}\{(\ln a_t - \ln a_{t-1}) + (\ln b_t - \ln b_{t-1}) + (\ln c_t - \ln c_{t-1})\}\end{aligned}$$

[14] In February 2013, Beverly 18, a movie theater in Champaign, IL, closed, and shortly after, Savoy 16, a neighborhood movie theater in Savoy, IL, opened a new I-Max theater, which provides an example of a negative spatial spillover effect at the disaggregated level data.

[15] That is, the root of $|I_R - A\, z| = 0$ from Eq. (3.5) falls outside the unit circle.

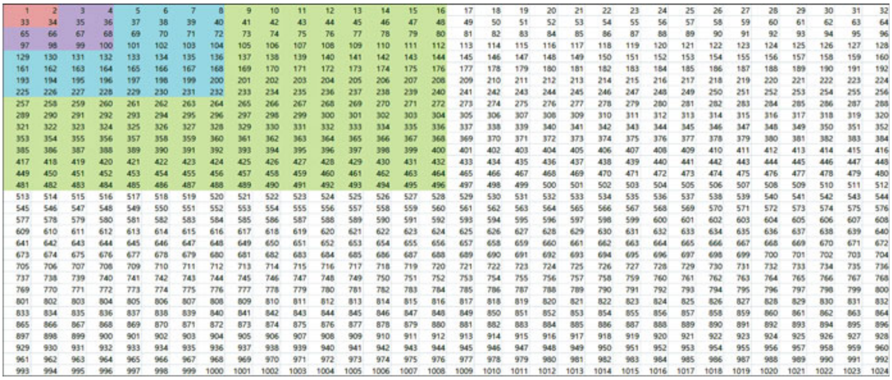


Fig. 3.2 Constructed regional economy

The true data generating process is defined at city level as shown in Eq. (3.5)[16]:

$$Y_t = \mathrm{A}Y_{t-1} + \varepsilon_t, t = 1, \ldots, T \tag{3.5}$$

where for a single-level structure,

[16]A more general version of this kind of structure can be expressed as equation (∗):

$$Y_t = \sum\nolimits_{i=1}^{p} A_i Y_{t-i} + \varepsilon_t,\ \ t = 1, \ldots, T \tag{$*$}$$

where

$Y_t = \left(y_t^1, y_t^2, \ldots y_t^R\right)'$ is an $R \times 1$ dependent variable,
$\{A_i | i = 1, .., p\}$ are $R \times R$ coefficient matrices, $E(\varepsilon_t) = 0, E(\varepsilon_t {\varepsilon_t}') = \Sigma$ (a positive definite coviariance matrix), $E(\varepsilon_t {\varepsilon_{t'}}') = 0 \forall t \neq t'$

For stationarity, all roots of $\left|I_R - \sum_{i=1}^{p} A_i z^i\right| = 0$ fall outside the unit circle. Also, (∗) can be expressed as VMA form as in equation (∗∗):

$$Y_t = \sum\nolimits_{i=1}^{\infty} \Phi_i \varepsilon_{t-i} \tag{$**$}$$

where

$\Phi_i = \sum_{j=1}^{p} A_j \Phi_{i-j}, i = 1,2, \ldots, \Phi_0 = I_R$, and $\Phi_i = 0\ \forall\ i < 0$. Since in this general case where the error term structure is not diagonal, the time profile of the shock affects the FEVD and CIRF; thus in our case, the error term structure is set to be diagonal for simplicity.

$$Y_t = \left(y_t^1, y_t^2, \dots y_t^{1024}\right)', \quad A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,1024} \\ \vdots & \ddots & \vdots \\ a_{1024,1} & \cdots & a_{1024,1024} \end{pmatrix} \text{ and}$$
$$\varepsilon_t = \left(\varepsilon_t^1, \varepsilon_t^2, \dots \varepsilon_t^{1024}\right) \sim N(0, \sigma^2 I_{1024}), \sigma^2 = 1,$$

and for a multi-level structure,

$$Y_t = \left(y_t^C, y_t^1, y_t^2, \dots y_t^{1024}\right)', \quad A = \begin{pmatrix} a_{C,C} & a_{C,1} & \cdots & a_{C,1024} \\ a_{1,C} & a_{1,1} & \cdots & a_{1,1024} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1024,C} & a_{1024,1} & \cdots & a_{1024,1024} \end{pmatrix} \text{ and}$$
$$\varepsilon_t = \left(\varepsilon_t^C, \varepsilon_t^1, \varepsilon_t^2, \dots \varepsilon_t^{1024}\right) \sim N(0, \sigma^2 I_{1024}), \sigma^2 = 1.$$

$\left\{y_t^r \mid r = 1, \dots, 1024\right\}$ are dependent variables representing the regional economic behaviors, whereas y_t^C is a region common factor. As noted earlier, the region common factor is little affected by local shocks by assumption; the elements $\{a_{C,i} \mid i = 1, \dots, 1024\}$ are set to be zero in the coefficient matrix. One might claim that this exogeneity assumption is rather extreme because, in reality, some spatial units or events can be powerful enough to affect the region common behavior. However, in the case where the region common factor is endogenous, the region common component is equivalent to just adding another regional unit in Eq. (3.5) with a single-level structure. Thus, by looking at the generated FEVDs with larger coefficients assigned to the regional units, we can easily conjecture that the results are drawn from an endogenous region common factor because the results should lie somewhere between the multi-level structure model and the single-level structure model.

Another critique can be raised regarding the diagonal error term structure. This error term structure does not allow the contemporaneous propagation of the shocks. This can be remedied if we disregard the $t = 0$. For example, if we observe the propagation of regional shocks at $t = 1$, some of the effects are already propagated into neighboring regions. Thus, whether the error terms are diagonal or not does not significantly change the results we are trying to draw from the exercise.[17]

The data generating process defines how the constructed regions interact through time and space. In this constructed regional economy, every regional interaction is determined by the coefficient matrix A such that growth in region j at time $t - 1$ will induce $a_{i,j}$ growth in region i at time t (neighborhood effect). Note that the above process is defined using demeaned variables; thus at the steady state, the growth rate of every region will grow at the historical average, and at the steady state, the value converges to zero ($Y_{\text{steady state}} = 0$). Thus, a typical local policy shock occurred in the center of the regional economy will exhibit a cumulative impulse response function

[17]The generalized versions of FEVDs and CIRFs with different spatiotemporal scale considering the contemporaneous shocks are also provided in the footnotes of the next subsection, in case the readers of this chapter wants to conduct a simulation with more complex regional economic system.

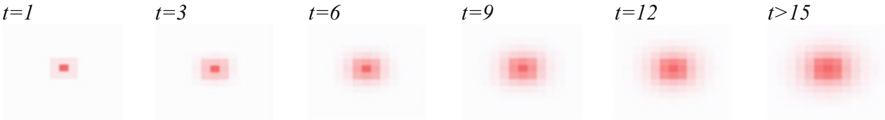


Fig. 3.3 A typical CIRF of spatially stationary process

(CIRF) graph such as in Fig. 3.3.[18] When a local shock is occurred at $t = 0$ at the center of the regional economy, it will propagate into its neighboring regions at $t > 0$. The marginal propagation effect extends forever. However, the amount of the effect eventually converges to zero, so that at $t > 15$, the CIRF remains almost the same shape.

3.3.2 Some Points to Note Before Deriving the Aggregated Form of FEVD[19]

It is neither necessary nor useful to derive the data generating process at the aggregated level in our analysis of the spillover effects. Since the purpose of the exercise is to see what we expect to observe at the aggregated level in terms of spillovers, where the amount of spillovers are defined by FEVD in this chapter, we can look directly at the aggregated form of FEVD instead of deriving the aggregated level data generating process.

The derivation of the data generating process for the temporally aggregated observations is relatively easier than that for the spatially aggregated observations. In our case, where there is only one lagged dependent variable and an i.i.d. error term with unit variance, the VAR(1) form can be preserved. For an $n-$ period aggregation, i.e., $Y_\tau = \sum_{i=0}^{n-1} Y_{t-i}$, Eq. (3.8) simply transforms into $Y_\tau = A^n Y_{\tau-1} + \epsilon_\tau$, where $\epsilon_\tau = \sum_{i=0}^{n-1} \left(A^i \sum_{j=0}^{n-1} \varepsilon_{t-i-j} \right)$. However, in practice, since the aggregated error term, ϵ_τ, is the superposition of n-multivariate normal distributions and if we assume approximate normality on the error term of the aggregated form, then it suffers a problem that we are automatically assuming the effects of the innovations at all n-subperiods are the same within one time period at the aggregated level. In this case, the economic translation of FEVD or CIRF using the aggregated form of data generating process becomes different from the original disaggregated data generating process. In more detail, for the case of FEVD, an element $\theta_{rs}(h)$ in FEVD is defined as the proportion of the h-step ahead forecast error variance of region r that is accounted for by the innovations in region s. Since $\epsilon_\tau = \sum_{i=0}^{n-1} \left(A^i \sum_{j=0}^{n-1} \varepsilon_{t-i-j} \right)$, innovations at subperiods $t-i$ in the aggregated error term are associated with t-i

[18]A graphical example of a CIRF of a spatially non-stationary process is provided in Appendix 4.
[19]Temporally/spatially aggregated forms of CIRFs are provided in Appendix 5.

different coefficients.[20] Thus, when we derive FEVD using the aggregated form of the data generating process, we are ignoring these differences. For example, an innovation in January and one in March are treated as the same when we derive h-quarter ahead forecasting error variance decomposition with the assumption of the normality of ϵ_τ.

Similar problem occurs when we derive the data generating process for the spatially aggregated observations. Following Arbia and Petrarca (2011), suppose G is an $S \times R$ aggregation matrix where $\frac{R}{S} = n$ (each aggregated level spatial unit contains the same n number of disaggregated level units).[21] Then, an aggregated form of the data generating process can be expressed as $Y_t^* = BY_{t-1}^* + \varepsilon_t^*$ where $Y_t^* = GY_t$, $B = GAG'(GG')^{-1}$, and $\varepsilon_t^* = G\varepsilon_t$. If we assume approximate normality, the aggregated level error term will be expressed as $\varepsilon_t^* \sim (0, n\sigma^2 I)$, as in Arbia and Petrarca (2011). However, as a matter of fact, ε_t^* is also a superposition of multivariate normal distributions, thus suffering from the same problem as in the temporal aggregation case. In other words, for the example in our constructed regional economic system, the approximate normality assumption will treat a unit shock on peripheral region such as city #1 the same as a unit shock on the region located closer to the center, such as city #34.[22]

3.3.3 *Derivation of the Aggregated Form of FEVD*

Since our objective is to see what we will observe at the aggregated level, we do not have to numerically derive the aggregated form of the data generating process. Instead, we can simulate some coefficient structure and visualize FEVD as the same fashion as in Sect. 3.3. In this manner, we can avoid the normality assumption problems mentioned previously.

In order to visualize what will happen in FEVD, another assumption that the shocks at the disaggregated level are distributed evenly across the initial aggregated period has been made.[23] Thus, the economic meaning of FEVD at the aggregated

[20] For example, $\epsilon_\tau = \cdots + (I + A + A^2)\varepsilon_{t-2} + \cdots$. This hinders researchers from performing efficient Monte Carlo type of simulation study.

[21] For example, when aggregating two units into one aggregated level unit ($n = 2$) where there are four spatial units, $G = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$.

[22] More specifically, when a regional shock spills over to region-sharing borders (rook contiguous), then region #1 spills over to regions #2 and #33, whereas region #34 spills over to regions #2, #33, #35, and #66.

[23] For example, aggregating at the quarterly interval, FEVDs are calculated assuming that the same amounts of shocks are given for the first 3 months. One can also simulate and visualize the case that shocks are unevenly distributed across within an aggregated time period, but since there are infinitely many cases of uneven distributions, and since even distribution is representative, only evenly distributed case is visualized here.

level, for example, an element $\theta_{rs}(h)$ in FEVD, is defined as the proportion of the h-step ahead forecast error variance of region r that is accounted for by the same amount of innovations in region s at time $t = 0, 1, \ldots, n - 1$ ($\in \tau = 0$).

Using this assumption, since the disaggregated level FEVD is as in Eq. (3.6), the aggregated level FEVD can be derived as in Eq. (3.7)[24]:

$$\theta_{rs}(h) = \frac{\sum_{l=0}^{h} \left(e_r' A^l e_s\right)^2}{\sum_{l=0}^{h} \left(e_r' A^l \left(A^l\right)' e_r\right)} \tag{3.6}$$

where

e_r is a $R \times 1$ selection vector (where the rth element $= 1$ with zeros elsewhere)

$$\theta_{rs}(H) = \frac{\sum_{m=0}^{n-1} \sum_{l=m}^{h+m} \left(e_r' A^l e_s\right)^2}{\sum_{m=0}^{n-1} \sum_{l=m}^{h+m} \left(e_r' A^l \left(A^l\right)' e_r\right)} \tag{3.7}$$

where

$H \ni \{h, h + 1, \ldots, h + n - 1\}$

For example, when aggregating monthly data into quarterly data, FEVD will be looking at the 12th month for the monthly model, and 12–14 months for the quarterly model, as shown in Fig. 3.4. In this case, the 1-year ahead FEVD becomes a weighted average of 12–14 months ahead of the monthly level FEVD with three consecutive monthly shocks at the first 3 months.

Similarly, from Eq. (3.6), a spatially aggregated version of FEVD can be expressed as in Eq. (3.8)[25]:

$$\theta_{RS}(h) = \frac{\sum_{l=0}^{h} \left(e_R' A^l e_S\right)^2}{\sum_{l=0}^{h} \left(e_R' A^l \left(A^l\right)' e_R\right)} \tag{3.8}$$

where R and S are aggregated spatial units; thus e_R and e_S are $R \times 1$ selection vectors where any rth and sth elements that belong to R and S, respectively, are ones and zeros elsewhere.

---

[24]For more general case, as in equation (∗), FEVD can be derived as $\theta_{rs}(h) = \frac{\sum_{l=0}^{h} \left(e_r' \Phi_l P e_s\right)^2}{\sum_{l=0}^{h} \left(e_r' \Phi_l \Sigma \Phi_l' e_r\right)}$. Thus, a temporally aggregated version of FEVD should be $\theta_{rs}(H) = \frac{\sum_{m=0}^{n-1} \sum_{l=m}^{h+m} \left(e_r' \Phi_l P e_s\right)^2}{\sum_{m=0}^{n-1} \sum_{l=m}^{h+m} \left(e_r' \Phi_l \Sigma \Phi_l' e_r\right)}$.

[25]Likewise, more generalized version of the spatially aggregated version of FEVD can be expressed as $\theta_{RS}(h) = \frac{\sum_{l=0}^{h} \left(e_R' \Phi_l P e_S\right)^2}{\sum_{l=0}^{h} \left(e_R' \Phi_l \Sigma \Phi_l' e_R\right)}$.

Monthly Data

shock 12-step ahead = 1 year ahead

↓												↑			
t=0	1	2	3	4	5	6	7	8	9	10	11	**12**	13	14	15

Quarterly Data

shock 4-step ahead = 1 year ahead

↓	↓	↓										↑	↑	↑	
t=0	**1**	**2**	3	4	5	6	7	8	9	10	11	**12**	**13**	**14**	15

Fig. 3.4 Conceptual comparison of 1-year ahead FEVD of monthly and quarterly data
∗Autoregressive coefficient for common factor = 0.2, autoregressive coefficient for regional factor = 0.1, neighborhood effect for multi-level structure = 0.22

For the temporal aggregation, the disaggregated level (first level) observations are aggregated with 2 (second level)–12 (twelfth level) periods of time. For the spatial aggregation, the disaggregated level (first level) observations are aggregated by four regional units at each level, up to the fifth level where the number of regional units is only four.

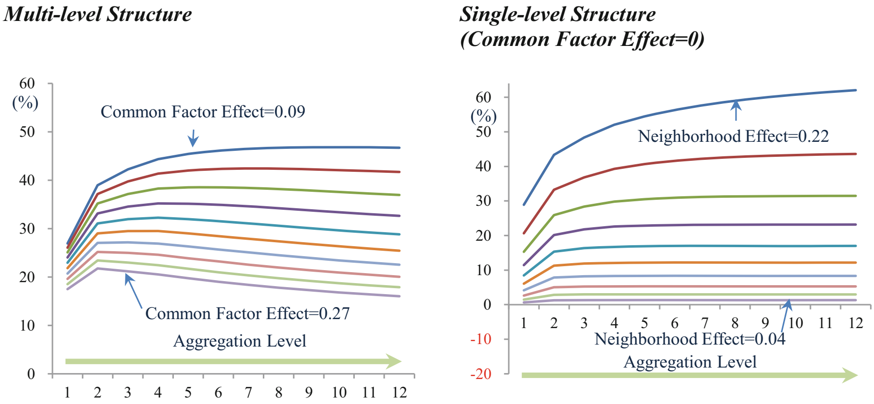
3.3.4 *Change of the Neighborhood Portion with Aggregation: FEVD Results*

When we decompose the sources of the variances of regional activities into neighborhood and its own (and region common for the multi-level structure) innovations, on average, the neighborhood portion decreases with larger scale of temporal/spatial observations. In the temporal aggregation case, however, when the effect from the region common factor is relatively small, the neighborhood portion increases as a result of a certain level of temporal aggregation, and for the extreme case where the region common factor effect is zero (single-level structure regional economic system), the neighborhood portion monotonically increases.

The FEVDs with various values of coefficients are calculated. In Eq. (3.5), the autoregressive coefficient for region common factor is set to be 0.2 ($a_{C,\,C} = 0.2$), and the effect from the region common factor is set to be from 0.09 to 0.27 ($\{a_{i,\,C} = 0.09\text{\textasciitilde}0.27 \mid i = 1, \ldots, 1024\}$). The autoregressive coefficient of regions own is set to be 0.1 ($\{a_{i,\,i} = 0.1 \mid i = 1, \ldots, 1024\}$). Each regional unit is assumed to be affected by its neighbors sharing common borders (rook contiguous), and various values of the effects are tried.[26]

The FEVDs of individual regional units vary a great deal depending on their locations; thus, it is not appropriate to try to find a specific pattern related to the

[26]Other values of autoregressive coefficient for region common factor, effect from the region common factor, and autoregressive coefficient of regions own are also tried, but not presented here, because the conclusions drawn from the results are same.



* Autoregressive Coefficient for Common Factor = 0.2, Autoregressive Coefficient for Regional Factor = 0.1, Neighborhood Effect for multi-level structure=0.22

Fig. 3.5 1-year ahead FEVD—neighborhood portion with different temporal aggregation level ∗Autoregressive coefficient for common factor = 0.2, autoregressive coefficient for regional factor = 0.1, neighborhood effect for multi-level structure = 0.22

change of the scale of units by looking at the individual level regional units.[27] However, when we average out our observations, we can see a clear pattern. The results for the temporal aggregations are presented in Fig. 3.5.[28]

Overall, the decrease of the neighborhood portion in the temporal aggregation case depends on the relative importance of the region common factor. For example, as shown in Fig. 3.5, when the effect from the common factor is 0.09 ($a_{i,\,C} = 0.09$), the neighborhood portion decreases after the eleventh level of aggregation, but when the effect from the common factor is 0.21 ($a_{i,\,C} = 0.21$), the neighborhood portion decreases after the third level of temporal aggregation. In the extreme case where the region common factor effect is zero, i.e., in a single-level structure model, the portion of the neighborhood innovation becomes larger with the temporal aggregations.[29] This result occurs because the spillover effects are propagating across multiple regional units and grow rapidly over time, whereas the autoregressive effect ($a_{i,\,i}$) remains in a single region, thus growing at a slower rate. However, in a multi-level structure economy, since the region common innovation and the neighborhood innovations are both propagating across multiple regions, depending on their relative importance, the neighborhood portion decreases after a certain level of aggregation. In sum, a larger value of the region common factor loading induced a more rapid

[27]As it is already shown in the real-world data example, the portion of neighborhood effect varies depending on the region. For our constructed regional economy, the variance of a region located at the border has larger portion of its own innovations.

[28]The results do not change much when the neighborhood coefficients are negative. Appendix 6 provides the results with negative neighborhood effects.

[29]This phenomenon also appears when we assign the autoregressive coefficients larger value such as $a_{i,\,i} = 0.8$.

decrease in the neighborhood portion. Thus, in a sense, neighborhood effects and common factor effects are two competing sources of change in the evaluation of the regional economic system. With more temporally aggregated observations, the region common factor loadings become more important, and we should expect less spatial dependency.

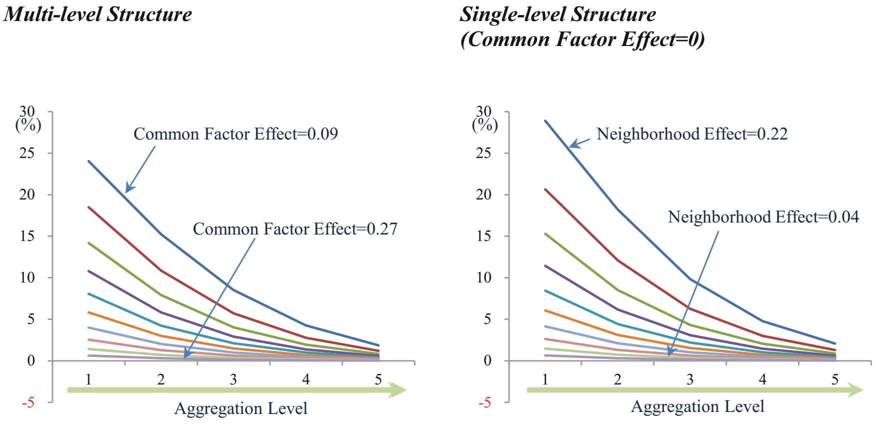
This trend over temporal aggregation concords with the average trends found in the multi-level estimation results from the real-world data in the previous section. In the real-world data, the portion of neighborhood innovation decreased in the exercise, implying that the region common factor occupies a larger portion of the regional variances.

Contrary to the temporal aggregations, spatial aggregation shows a monotonic decrease of the neighborhood portion regardless of whether the regional economy is a multi-level structure or a single-level structure. As shown in Fig. 3.6, regardless of our choices of the values for the common factor effects and neighborhood effects, the amount of spillover decreases monotonically.[30]

This monotonic decrease with the level of spatial aggregation is mainly due to the fact that our regional economic system is spatially stationary. Since all the innovations are processed in a spatially stationary system, an effect of a local shock fades away with distance. Thus, a shock once regarded as a *neighborhood innovation* can become an *own innovation* at a more aggregated level. In other words, unlike the temporal aggregation exercise where the relative portions of sources of innovations depend on the speed of their propagation across regional units, the spatial aggregation exercise is comparable to watching the regional economic system with lenses with different focal lengths at the same time spot, as described in Fig. 3.7.

As shown in Fig. 3.7, at a disaggregated level, we see the cumulative response of a local shock at the most detailed precision; thus the response around the innovation is categorized as a neighborhood effect. However, at the fourth-level aggregation, those neighborhood effects are mostly trapped inside the aggregated spatial unit; thus the neighborhood effect almost disappears. Relating this idea with FEVDs, with the assumption that the innovations are equally generated across the regions, only those local innovations located near the border of the aggregated spatial unit can penetrate into the neighboring aggregated spatial units. In other words, as long as the regional economic system is spatially stationary, spatial aggregation will result in smaller spillover effects.

[30]The results do not change much when the neighborhood coefficients are negative. Appendix 7 provides the results.



* Autoregressive Coefficient for Common Factor = 0.2, Autoregressive Coefficient for Regional Factor = 0.1, Neighborhood Effect for multi-level structure=0.22

Fig. 3.6 Twelve-step ahead FEVD—neighborhood portion with spatial aggregation level

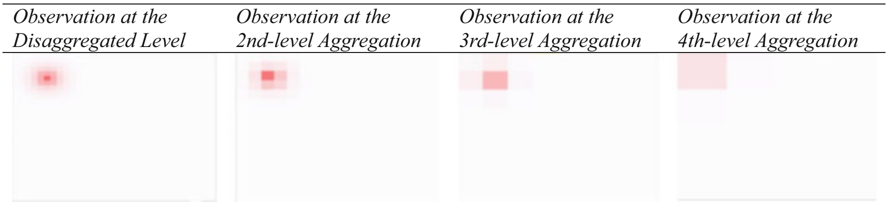


Fig. 3.7 Example of spatial CIRFs to a local shock with different levels of aggregations

3.4 Conclusion

The FEVD results on different levels of spatial/temporal aggregations are in agreement with Gosset's prediction that a larger scale of unit will reduce the spatial correlation. The results are revealed that, in temporal aggregation case, the amount of spillover effect can increase or decrease depending on the degree of the force of the region common factor. However, it is possible to conclude that we will observe less spillovers with larger spatial scale of units.

One caution should be made regarding the exercise of using a different observation scale. With the aggregation in terms of both space and time, the error term structure becomes a superposition of Gaussian normal distributions. Of course if the aggregated error term is a composition of considerable amounts of normal error terms, it can be approximated into a single normal error term, but in a situation where just several error terms are combined, which is in our case, this approximation can lead to different implications compared to the original data generating process. Thus, it is not appropriate to use Monte Carlo type of simulations when dealing with aggregated forms of data generating processes since the aggregated form does not reflect the changes of the error term structures with different levels of aggregations, inducing incorrect implications about the amounts of the spillovers. It would be appropriate to numerically calculate the aggregated form of FEVD and measure the percentage portion of the spillovers.

The characters of the constructed regional economic system in this chapter is that local shocks are transmitted through a VAR form of a coefficient matrix, and the coefficient matrix is constructed such that a local shock transmits to its immediate neighbor in the next period, i.e., it is a rook contiguity matrix. However, as long as the transmission channel of the local shock is dependent on geographical distance, so that the coefficient matrix defining the neighborhood effect has entries such as inverse distance or queen contiguity matrix, the conclusions are the same. The inclusion of negative entries in the coefficient matrix does not change the conclusions, either. This is more or less due to the necessary assumptions that the regional economic system is spatially stationary and that the neighborhood effects are geographically constrained within neighboring units. In other words, since the effect of an innovation fades away quickly with temporal/spatial distance, and the spatial aggregation binds regional units located close to each other, the spillover parts of innovation should be reduced with the larger scale of spatial units.

There are some limitations with our FEVD analysis on the constructed regional economic system. If the neighborhood effect is determined not by geographical closeness but by, for example, trade linkage, the spatial aggregation based on geographical location will not necessarily produce an inverse relationship between the neighborhood portion of FEVD and the level of spatial aggregation. Nevertheless, the exercise performed in this chapter shows less neighborhood dependency with larger scale of units. In many cases, we can expect smaller spillover effects with larger spatial scale because most human activities are physically constrained by their geographical locations, as shown in our real-world data exercise. For example, if we are working on the empirical studies using country data, it might be misleading to consider spillover effects between countries unless there are some trade linkages that can be observed explicitly.

Additional limitations relate to the relative sizes of regional units, the regional differences in the region common factor loadings, closed economy assumptions, and so forth, but how the relaxation of these strong assumptions will change the conclusions of this chapter remains a challenge for future analysis.

References

Arbia G (1989) Spatial data configuration in statistical analysis of regional economic and related problems. Kluwer Academic, Dordrecht

Arbia G, Petrarca F (2011) Effects of MAUP on spatial econometric models. Lett Spat Resour Sci 4:173–185

Bai J, Wang P (2012) Identification and estimation of dynamic factor models. Discussion Paper No.:1112-06, Department of Economics, Columbia University, New York, NY 10027

Barker T, Hashem Pesaran M (1990) Disaggregation in econometric modeling – an introduction. In: Barker T, Pesaran MH (eds) Disaggregation in econometric modelling. Routledge, London/New York
Brewer KRW (1973) Some consequences of temporal aggregation and systematic sampling for ARMA and ARMAX models. J Econ 1:133–154
Chung S, Hewings GJD (2015) Competitive and complementary relationship between regional economies: a study of the great Lake states. Spat Econ Anal. https://doi.org/10.1080/17421772.2015.1027252
Clark WAV, Avery KL (1976) The effects of data aggregation in statistical analysis. Geogr Anal 8:428–438
Corrado L, Fingleton B (2011) Multilevel modelling with spatial effects. SIRE discussion paper, SIRE-DP-2011-13
Dusek T (2004) Spatially aggregated data and variables in empirical analysis and model building for economics. Cybergeo: Revue européenne de géographie, 2004, No.285, p 21. www.cybergeo.presse.fr
Fotheringham AS, Wong DWS (1991) The modifiable areal unit problem in multivariate statistical analysis. Environ Plan A 23:1025–1044
Gehlke CE, Biehl K (1934) Certain effects of grouping upon the size of the correlation coefficient in census tract material. J Am Stat Assoc 29:169–170
Gelfand AE, Diggle PJ, Fuentes M, Guttorp P (eds) (2010) Handbook of spatial statistics. CRC Press, Boca Raton
Griffith DA, Wong DWS, Whitfield T (2003) Exploring relationships between the global and regional measures of spatial autocorrelation. J Reg Sci 43:683–710
Haitovsky Y (1973) Regression estimation from grouped observations. Griffin, London
Marcellino M (1999) Some consequences of temporal aggregation in empirical analysis. J Bus Econ Stat 17:129–136
McCullagh P, Clifford D (2006) Evidence for conformal invariance of crop yields. Proc R Soc A Math Phys Eng Sci 462(2071):2119–2143
Okabe A, Tagashira N (1996) Spatial aggregation bias in a regression model containing a distance variable. Geogr Syst 3:77–99
Openshaw S, Taylor PJ (1979) A million or so correlation coefficients: three experiments on the modifiable areal unit problem. Stat Appl Spatial Sci 21:127–144
Pearson ES (1990) Student: a statistical biography of William Sealy Gosset. In: Plackett RL, Barnard GA (eds) Oxford University Press, Oxford
Prais SJ, Aitchison J (1954) The grouping of observations in regression analysis. Rev Int Stat Inst 22:1–22
Smith HF (1938) An empirical law describing heterogeneity in the yields of agricultural crops. J Agric Sci 28:1–23
Tagashira N, Okabe A (2002) The modifiable areal unit problem, in a regression model whose independent variable is a distance from a predetermined point. Geogr Anal 34:1–20
Wei WWS (1981) Effect of systematic sampling on ARIMA models. Commun Stat Theory Methods 10:2389–2398
Weiss AA (1984) Systematic sampling and temporal aggregation in time series models. J Econ 26:271–281
Whittle P (1956) On the variation of yield variance with plot size. Biometrika 43:337–343
Whittle P (1962) Topographic correlation, power-law covariance functions, and diffusion. Biometrika 49:305–314
Wong DWS (2004) The modifiable areal unit problem (MAUP). In: Janelle DG, Warf B, Hansen K (eds) Worldminds: geographical perspectives on 100 problems. Springer, Dordrecht, pp 571–575

Chapter 4
Population Characterization in Location Modeling: Alternatives, Impacts, and Insights

Daoqin Tong, Wangshu Mu, and Changfeng Li

Abstract Location analysis and modeling have been widely applied to support locational decisions for service provision. The general idea of such analysis has been for sited facilities to serve the demand of interest in an efficient and/or effective way. In many applications, service demand involves either general people or certain population groups in a region. Currently, population-based demand has been assessed mainly based on where people live, primarily using census population count data. This can be problematic given that people do not always stay at home or originate their trips from home. As a result, relying upon residential information may lead to an inaccurate evaluation of service demand in location modeling. This study investigates the impacts of alternative population characterizations on the classic *p*-median problem. A new model incorporating time-varying population distributions is introduced. An empirical study was conducted in three regions in Shanghai, China, where time-varying population distributions were derived using cell phone data. Analysis results show that solutions generated based on where people live can be far from the optimal that considers the temporal variability of population distributions. Discussion is provided on ways to remedy the issue.

Keywords Location modeling · Population · Demand

D. Tong (✉) · W. Mu
School of Geographical Sciences and Urban Planning, Arizona State University, Tempe, AZ, USA
e-mail: Daoqin.Tong@asu.edu

C. Li
China Academy of Urban Planning and Design, Beijing, China

R. S. Franklin (ed.), *Population, Place, and Spatial Interaction*, New Frontiers in Regional Science: Asian Perspectives 40,
https://doi.org/10.1007/978-981-13-9231-3_4

4.1 Introduction

Location analysis and modeling have been widely used to support locational decisions for various service provisions in both the public and private sectors. Applications include urban green space design (Zhang et al. 2017), emergency humanitarian logistics (Boonmee et al. 2017), health service development planning (Rahman and Smith 2000; Ahmadi-Javid et al. 2017), and farmers' market placement (Tong et al. 2012). The general idea of these models has been to locate facilities to serve the demand of interest in an efficient and/or effective way. While some models focus on ensuring a certain level of service with minimal resources, others try to achieve the maximal efficiency or equity given a limited budget.

In many location models, people often serve as the demand for the intended services. Such services include healthcare (Murawski and Church 2009; Meskarian et al. 2017), public transportation (Wu and Murray 2005), cell phone signal coverage (Akella et al. 2010), emergency responses (Marianov 2017), and disaster relief good planning (Widener and Horner 2011; Chen et al. 2013). In these studies, population has been mainly characterized based on where people live. Such information can be obtained from Census Bureau in certain aggregate form. For example, American Community Survey (ACS) provides population count data based on census data collection units. Using census data, Socioeconomic Data and Applications Center (SEDAC) provides population estimates at certain grid level (e.g., 1 km).

Characterizing demand based on where people live might be problematic in many real-world applications as people may need to be present at workplaces, move on roadways, and visit parks depending on the time of the day. Except for a few studies, relying on the static residential information for population representation may lead to an inaccurate evaluation of service demand in location modeling. It remains unknown to what extent such an inaccurate demand assessment affects the optimal solution and whether the impact varies with population distributions. Drawing on spatiotemporal cell phone data collected in Shanghai, China, this chapter aims to provide an investigation on these questions. The next section provides a literature review on population characterization in existing location analysis studies and the associated problems. This is followed by an introduction to a classic location model and a new model incorporating temporal variability of population. An empirical study is then conducted with the results presented. We conclude with some discussion and future research directions.

4.2 Background

In many location modeling studies, population-based demand is often approximated using census population count data given the data availability. The count data summarize the total population information in the associated data collection unit, such as census block group or tract. One common practice has been to aggregate these areal units into points each assigned with the corresponding population information. For example, in a study on the sexual health service provision, Meskarian et al. (2017) aggregated postcode-level population into a set of points representing the service demand. In seeking the best sites for locating mobile food markets, Widener et al. (2012) used block group centroids to represent the demand site for fresh food. While most studies directly rely on census data for population demand, some applications may need finer population information. To better characterize the spatial variation, Wu and Murray (2005) made a population interpolation analysis at the 30 $\times$ 30 m scale and applied the associated population estimates to determine the best modification strategy for transit service provision.

Using census population data as the proxy for service demand is equivalent to assuming people receive or originate their trips for the intended service at home. For certain services that people need to receive all the time such as cell phone signals and emergency responses, ensuring coverage of residential areas is insufficient. This is because in addition to home people frequently visit and stay at other important sites such as workplaces, schools, parks, etc. For example, the 2017 American Time Use Survey reported that workers on average spent about 8 h on an average weekday at work, and 83% of workers did some or all of their work at workplace. Also, people spend a significant amount of time traveling to their activity destinations. The 2017 National Household Survey found that on average, American drivers and passengers spent about an hour in a vehicle every day.

For services where people need to make their trips for, such as transit services, grocery stores, and medical care, current modeling results may only cover home-based trips by ensuring that the service provided is most convenient to homes. However, studies showed that people may also initiate their travel from other important locations such as school and workplace. Based on a survey, Mack and Tong (2015) reported that about 42% of the farmers' market trips originated from nonhome places. In general, nonhome-based trips have been found to account for over 30% of all daily trips (Mcguckin et al. 2005). Recognizing that workplaces may serve as important sites where people originate their trips from, several studies expanded the demand representation to also include employment in their location modeling. For example, in a transit stop removal study, Wu and Murray (2005) considered the amount of employees in each census block along with the census population for potential transit service. Similarly, in determining the best farmers' markets sites, Tong et al. (2012) considered workers at their workplaces for potential demand for farmers' markets. While in some applications, it is important to locate services/facilities close to where people live or work, a more general framework has been to capture the commute flows by locating facilities close to commute routes. The approach to incorporating commute-based trip chain has been used to site children day care centers (Hodgson 1981) and determine the location and operation time of farmers' markets (Tong et al. 2012).

In addition to home and workplace, people may be at different places during different times of the day for various purposes, and these may include a large number of non-commute chained trips. Based on a 2008 National Household Travel Survey add-on dataset, Li and Tong (2017) observed 78% non-commute chained trips. They then developed a facility location model to address a full spectrum of trip chaining by incorporating travelers' daily activity-travel program into service provision planning. According to their approach, people have the flexibility of visiting a facility from any activity site, including home, workplace, and other activity destinations, and along any trip made based on an individual's daily activity-travel program. Such an approach assumes the knowledge of an individual's daily activity-travel program, which can be very challenging due to the data availability issue.

To have a better understanding of the issue associated with population characterization in location modeling, this study analyzes the temporal variability of population distributions and examines how locational decisions may vary with alternative demand characterization. Different from detailed activity-travel data used in Li and Tong (2017), data involved in this study are relatively easier to obtain, especially considering the increasing availability of large geotagged data collected through cell phones and wearable devices. We also note the difference between people's daily movement and seasonal or long-term migration. For example, Ndiaye and Alfares (2008) provided a study on healthcare facility location where population groups migrated seasonally. In their study, during a season, population groups were fixed at home locations, and people's daily movement was not considered. We use the classic p-median problem (ReVelle et al. 2008) to demonstrate the nuances of alternative population characterization and investigate whether and how problem solutions may be impacted.

4.3 Methodology

4.3.1 PMP

The p-median problem (PMP) is one of the classic location problems that aims to site a number of facilities so that the overall demand-weighted travel distance/time to the closest facility is minimized. The problem was first introduced by Hakimi (1964, 1965) in a network context, where the optimal sites are called "medians" of the network. The PMP linear programming model was first provided by ReVelle and Swain (1970). Since then, the PMP has been widely studied in the literature. The problem has been applied to support cluster analysis (Klastorin 1985), transportation logistics (Pamučar et al. 2016), political redistricting (Hess et al. 1965), bike-sharing station planning (Park and Sohn 2017), and healthcare center siting (Jia et al. 2014).

Consider the following notation:

i: index of demand
j: index of candidate facility site
h_i: demand associated with i
d_{ij}: distance between i and j
p: the number of facilities to be sited

$$x_j = \begin{cases} 1 & \text{if candidate site } j \text{ is selected} \\ 0 & \text{otherwise} \end{cases}$$

$$y_{ij} = \begin{cases} 1 & \text{if demand } i \text{ is allocated to facility at } j \\ 0 & \text{otherwise} \end{cases}$$

The PMP can be formulated as

$$\text{Minimize } \sum_i \sum_j h_i d_{ij} y_{ij} \tag{4.1}$$

subject to

$$\sum_j y_{ij} = 1 \qquad \forall i \tag{4.2}$$

$$y_{ij} \le x_j \qquad \forall i, j \tag{4.3}$$

$$\sum_j x_j = p \tag{4.4}$$

$$x_j, y_{ij} \in \{0, 1\} \qquad \forall i, j \tag{4.5}$$

The PMP objective (4.1) minimizes the total demand-weighted travel distance. Constraints (4.2) require that each demand be assigned to one facility. Constraints (4.3) ensure that demand i is assigned to facility at j only when site j is selected for siting. Constraint (4.4) specifies p number of facilities to be sited. Constraints (4.5) impose binary conditions on decision variables. As specified in the original PMP, demand at i and h_i does not vary with time.

4.3.2 The PMP-Time-Varying Demand (PMP-TD)

As discussed previously, population distributions may vary with time (t), resulting in a time-varying demand distribution h_{it}. The corresponding problem will then become identifying the spatial configuration of p facilities so that they are most accessible, considering the temporal variability of demand. Consider the additional notation,

$$\mathrm{y}_{ijt} = \begin{cases} 1 & \text{if demand at } i \text{ is allocated to facility at } j \text{ during time } t \\ 0 & \text{otherwise} \end{cases}$$

d_{ijt}: distance between i and j during time t

The PMP-TD can be formulated as,

$$\text{Minimize} \sum_i \sum_j \sum_t h_{it} d_{ijt} y_{ijt} \tag{4.6}$$

subject to

$$\sum_j y_{ijt} = 1 \qquad \forall i, t \tag{4.7}$$

$$y_{ijt} \le x_j \qquad \forall i, j, t \tag{4.8}$$

$$\sum_j x_j = p \tag{4.9}$$

$$x_j, y_{ijt} \in \{0, 1\} \qquad \forall i, j, t \tag{4.10}$$

Objective (4.6) minimizes the overall demand-weighted travel across all times. Constraints (4.7) specify that demand at i during time t can be assigned to only one facility. Constraints (4.8) states that demand at i during time t can be assigned to facility at j only when site j is selected for siting. Constraint (4.9) is the same as constraint (4.4). Constraints (4.10) impose binary integer conditions on the decision variables.

We note here that given a spatial configuration of facilities, the assignment of demand i does not change with time. This is because in the PMP demand (h_{it}) at i is always assigned to the closest facility based on the minimal travel distance objective. That is, demand allocation only depends on the spatial distribution of i and j. Given that d_{ij} is fixed over time t ($d_{ij} = d_{ijt}$), y_{ijt} is always the same for a given pair of i and j. The time-independent nature of demand allocation y_{ijt} also makes constraints (4.7) irrelevant to time: given i and j, y_{ijt}=1 for one time period t' ensures satisfaction of constraints (4.7) for all other time periods. As a result, in the PMP-TD, y_{ijt} collapses into y_{ij}, and constraints (4.7, 4.8, 4.9, and 4.10) can be replaced by constraints (4.2, 4.3, 4.4, and 4.5). Therefore, the new problem involves a new assessment of demand $\sum_t h_{it}$ at demand site i. The new demand is a sum of the demand at i across all times. Objective (4.6) then becomes $\sum_i \sum_j \sum_t h_{it} d_{ij} y_{ij}$ (11)

4.4 Empirical Study

We will use a case study to demonstrate how to incorporate time-varying population distributions into the PMP. We will also compare whether and how the solutions based on time-varying demand differ from those obtained based on where people live. The case study consisted of three regions in Shanghai, China (also see Fig. 4.1). The first region contains Lujiazui and its surrounding neighborhood communities with an overall area of 48 km^2. This region serves as one of central business districts (CBDs) in Shanghai and is known as one of the most important financial districts in China. The second region is composed of the southern part of Yangpu district. This region is primarily residential area with a size of 30 km^2. The third region is located in Zhangjiang Town with an area of 50.4 km^2. This region contains a major

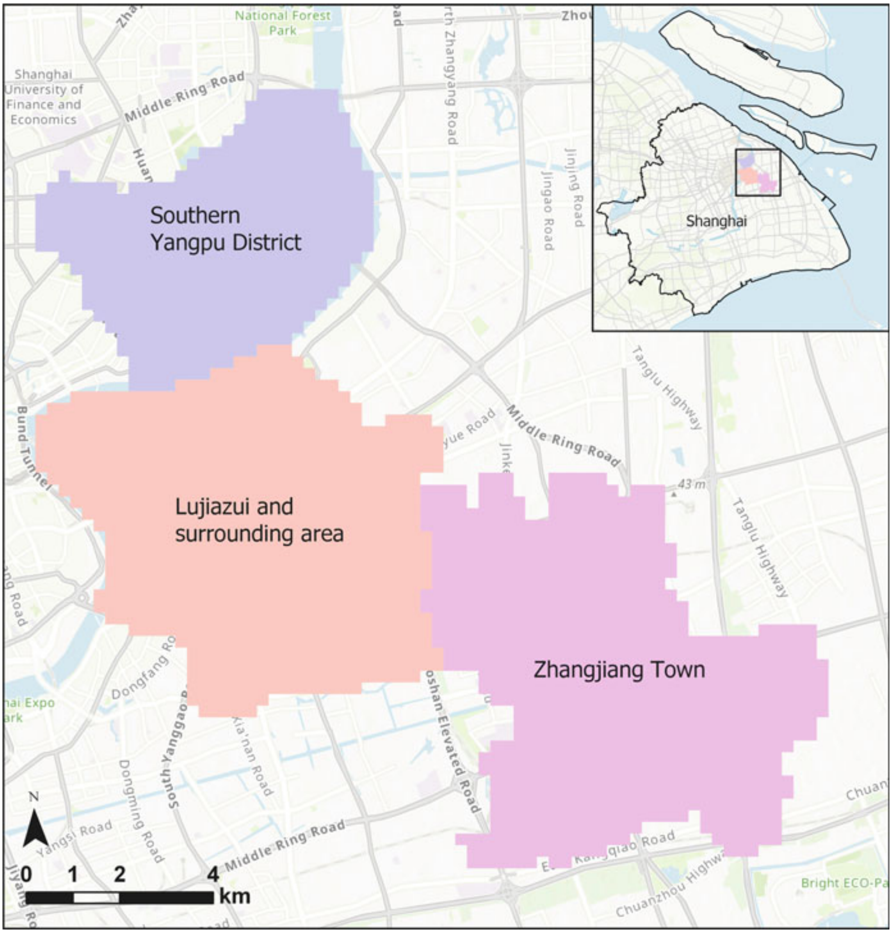


Fig. 4.1. The study area

technology park hosting many IT companies in the northwest and some residential areas in the southeast.

Spatiotemporal distributions of the population in the three regions were derived based on cell phone data provided by China Unicom. The data were collected for an entire week from Monday, November 20, 2017 to Sunday, November 26, 2017. Hourly population count was summarized using 250 $\times$ 250 meter grids. In each of the three regions, we assumed five facilities to be sited ($p = 5$). For each region, the PMP-TD was performed to obtain the optimal solution considering the temporal variation in the population distribution. Meanwhile, for each hour during weekdays and weekends, we obtained the PMP optimal solutions based on the population observed during that hour. For each of these solutions, we mapped the facility sites obtained and computed the overall travel using the time-varying demand. Such travel was then compared with the optimal travel obtained using the PMP-TD to compare the solution quality.

4.5 Results

4.5.1 Temporal Variability in the Population Distribution

Figure 4.2 shows the temporal variation of the population in the three regions during weekdays and weekends. The horizontal axis records the 24 h of a day; it starts from midnight (0) of the first day and ends before the midnight of the following day (23). On weekdays, compared to the midnight population, Lujiazui region gained a significant amount of population (58%) during the daytime, especially during the work hours (8 am–6 pm). This is not surprising given the CBD functionality of the region. In Zhangjiang Town, we also note significant population gain (56%) during weekday work hours. In contrast, Yangpu region had a stable population distribution with a daily population change of 8%. On weekends, while Yangpu and Zhangjiang had relatively consistent population throughout the day, Lujiazui attracted as much as 37% of people to this region as it also serves as an important tourist attraction site.

Figure 4.3 maps the spatial distribution of population gain/loss at 10 am compared with that at midnight for both weekdays and weekends. For each grid in a region, the population count at midnight is used as the baseline. Compared with the population at midnight, blue areas are places losing population at 10 am, whereas red areas correspond to locations gaining population. During weekday workday hours, Lujiazui has substantially more areas that gained population than areas that lost population (Fig. 4.3a). We notice that areas gaining population in Lujiazui were distributed extensively throughout the region. This is different from the pattern we observe in Zhangjiang (Fig. 4.3c). In Zhangjiang, areas gaining population are

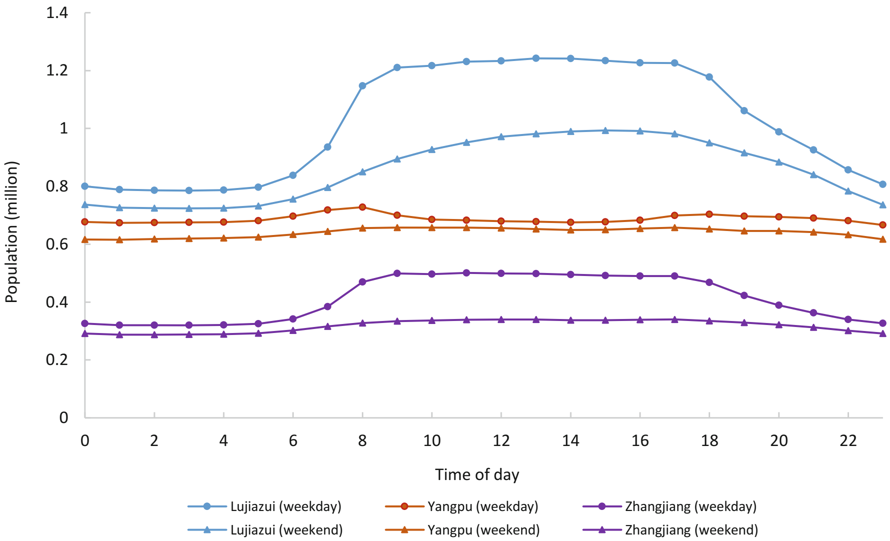


Fig. 4.2. Overall population change in the three regions

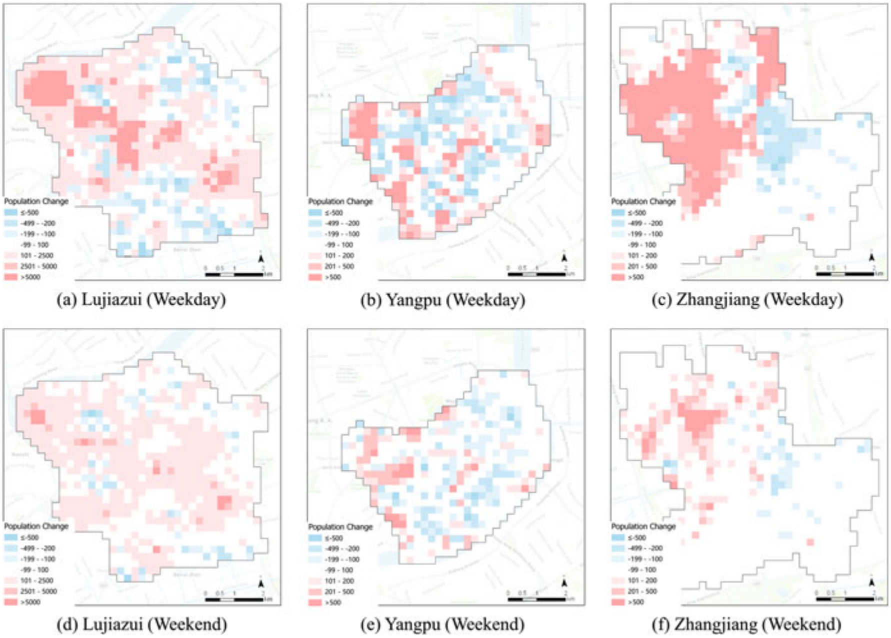


Fig. 4.3. Population gain/loss assessed at 10 am compared with midnight. (**a**) Lujiazui (weekday). (**b**) Yangpu (weekday). (**c**) Zhangjiang (weekday). (**d**) Lujiazui (weekend). (**e**) Yangpu (weekend). (**f**) Zhangjiang (weekend)

mainly clustered in the northwestern part of the town where the technology park is located, and areas losing population are concentrated in the residential areas east to the technology park. Different from Lujiazui and Zhangjiang, during weekday work hours, Yangpu has minimal areas gaining population, and these areas are highly dispersed in the region. During weekends, most of the areas gaining population in Lujiazui are similar to those during weekdays though with a smaller magnitude of gain. For Yangpu and Zhangjiang, much fewer areas gained population during weekends. For both weekdays and weekends, we notice that the magnitude of population gain in Lujiazui is much higher than the other two regions as reflected in the legend, indicating the ability of many zones in this region in attracting people during the daytime.

Figure 4.4 summarizes the average absolute population change (%) in a region for both weekdays and weekends. Similar to Fig. 4.3, we computed the change using the overall midnight population as the baseline. Population change in a grid at time t was computed using the percentage of population gain/loss at time t compared with the midnight population. Here, we did not differentiate population gain from loss as the focus is on the magnitude of change. Summarizing all grids in a region, the average population change was then calculated at time t. For all the three regions, population changes during weekdays were higher than those during weekends. Different from the overall population change shown in Fig. 4.2, Fig. 4.4 also captures people's

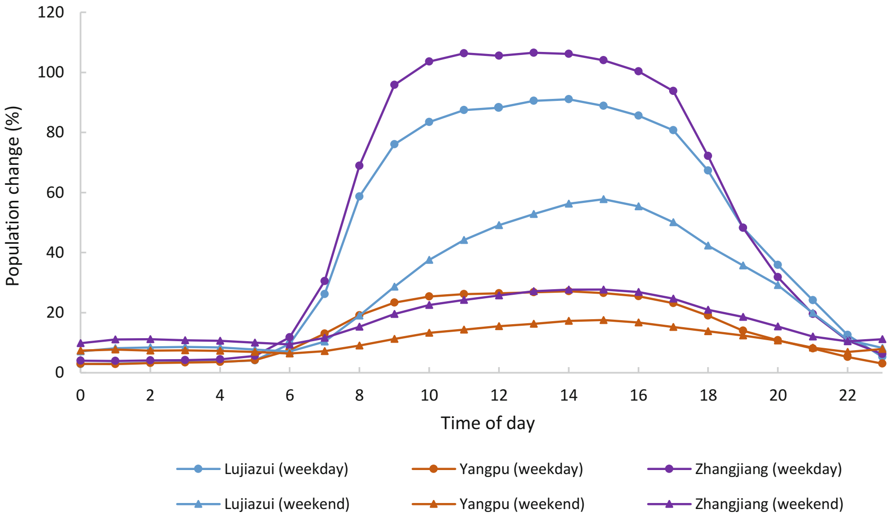


Fig. 4.4 Average absolute population change in the three regions

movement within a region. As shown in Fig. 4.4, in a general weekday, population changes were higher than the weekend population changes, which is also consistent with the overall population change in Fig. 4.2. However, when intraregional movement is considered, Fig. 4.4 gives different population change profiles. According to Fig. 4.4, on weekdays, Zhangjiang had the largest population change (106%), followed by Lujiazui (91%) and Yangpu (27%). This is different from the overall population change curve in Fig. 4.2, where Lujiazui had the highest overall population gain (58%), followed by Zhangjiang (56%) and Yangpu (8%). This suggests significant intraregional population exchange in Lujiazui and Zhangjiang. Unlike weekdays, Lujiazui had the highest average population change (58%) during weekends followed by Zhangjiang (28%) and Yangpu (18%).

4.5.2 Optimal Solution Comparison

We compared the PMP-TD solutions with the PMP solutions for three times (t), 0 am, 10 am, and 6 pm (Figs. 4.5, 4.6 and 4.7). While we separated the weekday and weekend solutions for $t = 10$ am and 6 pm, we used the average midnight population across the entire week to derive the solution for the PMP with $t = 0$ am, given that the population distribution at midnight did not vary much across days. Figures 4.5, 4.6, and 4.7 plot the facilities selected (black stars) and the associated allocation (black lines) of population to its nearest facility. For the PMP solutions (Figs. 4.5b–f, 4.6b–f, and 4.7b–f), the population at the corresponding time t is shown as the

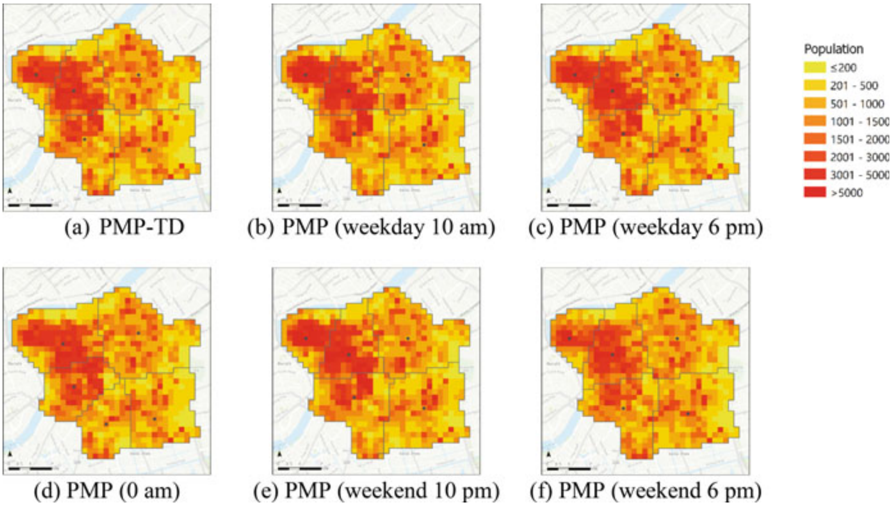


Fig. 4.5. Solution comparison in the Lujiazui region. (**a**) PMP-TD. (**b**) PMP (weekday 10 am). (**c**) PMP (weekday 6 pm). (**d**) PMP (0 am). (**e**) PMP (weekend 10 pm). (**f**) PMP (weekend 6 pm)

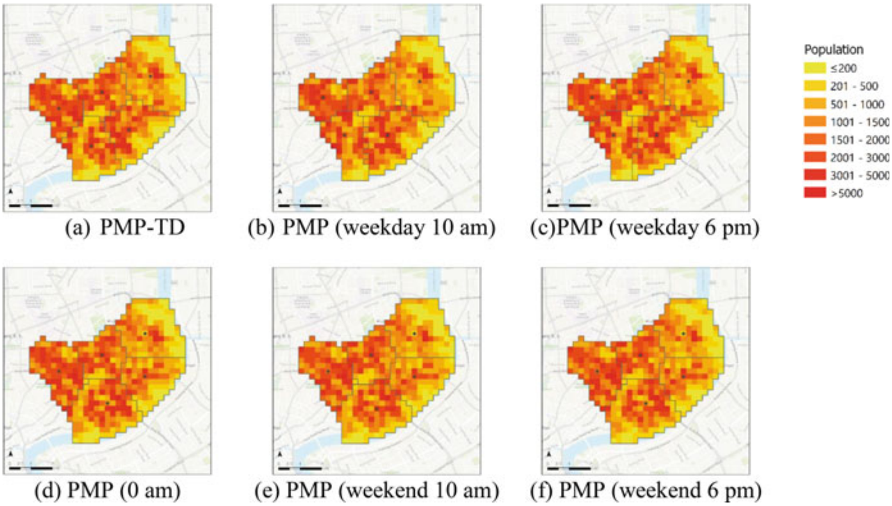


Fig. 4.6. Solution comparison in the Yangpu region. (**a**) PMP-TD. (**b**) PMP (weekday 10 am). (**c**) PMP (weekday 6 pm). (**d**) PMP (0 am). (**e**) PMP (weekend 10 am). (**f**) PMP (weekend 6 pm)

background. For the PMP-TD solutions, the overall average population incorporating the hourly variation across the entire week is mapped (Figs. 4.5a, 4.6a, and 4.7a).

For Lujiazui, the spatial configurations of sited facilities drawn based on the PMP during the daytime (e.g., $t = 10$ am and 8 pm) tend to resemble those given by the PMP-TD. This is as expected. As we show previously, the region gains a significant amount of population during the daytime (58%). The significantly higher demand in

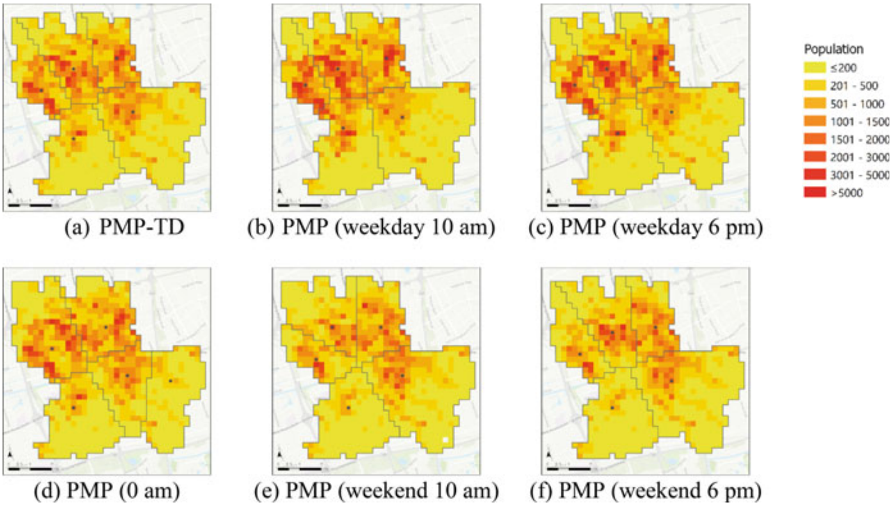


Fig. 4.7 Solution comparison in the Zhangjiang region. (**a**) PMP-TD. (**b**) PMP (weekday 10 am). (**c**) PMP (weekday 6 pm). (**d**) PMP (0 am). (**e**) PMP (weekend 10 am). (**f**) PMP (weekend 6 pm)

these hours lead to higher weights, given these hours in Objective (11), which eventually helps pull the PMP-TD optimal solution toward sites that best serve the population distribution during these hours. We also note that the population distribution during weekend daytime in this region is similar to that during weekday daytime with population concentrations in the northwestern part of the region. The PMP solutions during weekend daytime (Figs. 4.5e and f) are therefore similar to the PMP-TD solution. We notice significant difference in the spatial configuration of sited facilities when comparing the solution given by the PMP ($t = 0$) with that given by the PMP-TD. While both models prescribe three facilities to serve the western part of the region, the PMP-TD sites two facilities in the northwest, whereas the PMP ($t = 0$) locates only facility in that area.

In the Yangpu region, the solutions given by the PMP and PMP-TD are similar due to the overall small population change throughout the day. Only slight difference exists between the weekday and weekend solutions. We notice that weekend daytime PMP solutions are similar to the midnight PMP solutions. This is also to our anticipation, given that the weekend population distribution has the minimal temporal variation.

In Zhangjiang, we have a similar comparison observation to that in Lujiazui. The spatial configuration of the sited facilities using the PMP-TD (Fig. 4.7a) is very similar to that based on the PMP solutions during weekday daytime (Figs. 4.7b and c). This is because similar to Lujiazui, Zhangjiang gains substantial population (56%) during the weekday daytime. As we discussed previously, such an increase will result in higher weights in Objective (11) given to sites to better serve the weekday daytime population. The PMP solutions based on weekend and weekday daytime population are also similar except for the slight difference in the PMP weekend morning solution (Fig. 4.7e). The spatial configuration of the PMP midnight solution is found to be drastically different from that given by the PMP-TD solution: while the PMP-TD prescribes three facilities to serve the technology park area, the PMP locates only two facilities in the area (Fig. 4.7d).

We use the population distribution at midnight ($t = 0$) to approximate the census population. Assuming a spatial configuration of facilities sited using the PMP ($t = 0$), we computed the travel involved using the time-varying demand and compared it with the travel based on the PMP-TD solutions. Figure 4.8 shows the comparison for both weekdays and weekends. Here, positive/negative additional travel means the PMP solutions need more/less travel when compared with the PMP-TD solutions. In general, the PMP solutions involve more travel during daytime (e.g., 7 am–8 pm) and less travel at night (e.g., 9 pm–6 am). As for daytime, the PMP solutions require significantly more travel during weekdays than weekends with the highest weekday additional travel of 15.3% for Lujiazui and 15.2% for Zhangjiang, respectively, compared to highest weekend additional travel of 7.5% for Lujiazui and 3.3% for Zhangjiang, respectively. Combining weekdays and weekends, the PMP solutions require an additional daytime travel of 9.2% for Lujiazui, 0.8% for Yangpu, and 8.6% for Zhangjiang. As for nighttime, the PMP solutions give a travel reduction of 5% for Lujiazui and 3% for Zhangjiang. Summarizing the entire week, the additional travel brought about by the PMP is 4.1% for Lujiazui, 0.3% for Yangpu, and 4.3% for Zhangjiang.

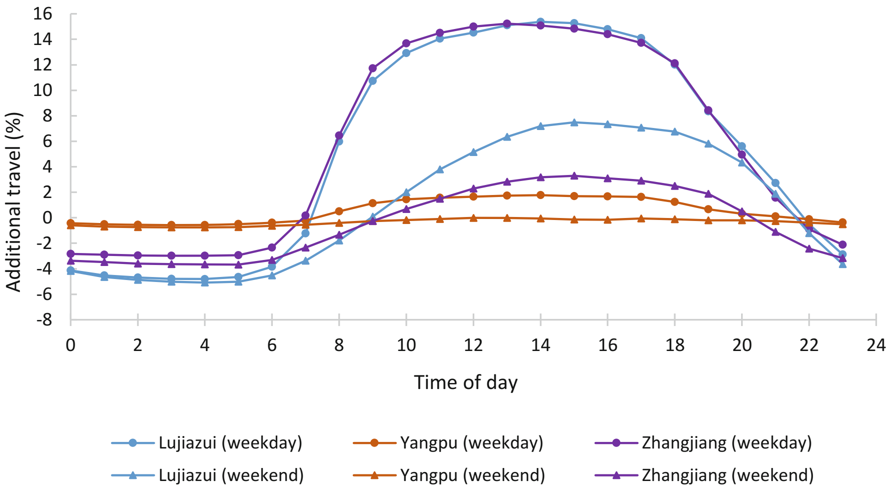


Fig. 4.8. Additional travel based on the PMP ($t = 0$) solutions

4.6 Discussion and Conclusion

In this study, we observe distinctly different population distributions by time of day (daytime vs. nighttime) and across days (weekdays vs. weekends) especially in Lujiazui and Zhangjiang. While many of existing location models focus on siting facilities based on where people live, our empirical study indicates that the solutions obtained using this approach could be very different from the optimal one when the temporal variation of population is considered. Using a classic location model as an example, we find that the existing approach may result in much worse solutions when the temporal variation of population is large. We note that for many public services that are only available during daytime, such as postal offices and public libraries, the existing approach may give even worse solutions. Although the existing approach appears to be more applicable during weekends, our empirical study shows that the nontrivial temporal variation of population in two of the three regions has led to significantly more travel needed by the existing approach.

Numerous studies have been developed to seek the optimal solutions to the PMP, especially for medium- and large-sized problems. Many of these approaches try to close the optimal gap at the magnitude of less than 1%. For example, Mu and Tong (2018) introduced a spatial-knowledge-enhanced Teitz and Bart (STB) algorithm for solving the PMP with an improvement of less than 1% for most test cases. Irawan and Salhi (2015) developed a PMP solution heuristic based on a demand aggregation strategy and reported an improvement of less than 0.5%. If the temporal variability of population is not considered, solutions provided by the PMP could be farther from the optimal when compared with heuristic approaches. As we show previously, compared to the PMP-TD, the PMP solutions have an overall optimality gap of 4% for Lujiazui and Zhangjiang and 0.3% for Yangpu.

In real-world applications, it is therefore worthwhile to examine the temporal distribution of the targeted population before the implementation of a location model. For an area where the population distribution does not vary much with time, as in the case of Yangpu, a direct application of the associated location model based on census population information might yield a solution that is not very far from the optimal one. If an area involves significant population change throughout the day, solely relying upon census population data can be very problematic. In this case, a better characterization of the population will be needed. As for the PMP, we show that the average population throughout the day will be appropriate. In this case, the 24-h average population data (e.g., LandScan data) could be used. However, whether such data are suitable for other location models remains unknown.

In this research, we incorporate the temporal variability of population into location modeling by discretizing the population distribution into a finite number of time periods. The population in each time period has an equal probability of being served by the sited facility. This assumption is more appropriate for certain applications, such as cell phone signal coverage and emergency services. However, it can be less appropriate for some other applications, such as grocery stores and dining places as people may be at work or have other constraints during a particular time period that may prevent them from using the service. In these applications, a better characterization of the population as potential demand will be needed.

Nowadays, big geospatial data have been widely collected through taxis, shared mobility applications such as Uber and Mobike, wearable devices, and social media platforms. The emergence of big data provides statistically sound samples with finer spatial and temporal resolutions. These data provides the opportunity to revisit some of the assumptions we make in many location models (Tong and Murray 2017). In this study, the temporal variation of population derived using cell phones allows us to examine the impact of the population assumption made in one classic location model. In addition, big data offer the opportunity for us to study individual-level mobility and travel activity, which will be helpful for a better characterization of service access. For example, the PMP assumes that people visit the closet facility, which might not always be true. Some people may chain a visit to a facility with other important trips even if the facility is far away. How to incorporate more complex travel behavior into location modeling points to another venue for future research.

References

Ahmadi-Javid A, Seyedi P, Syam SS (2017) A survey of healthcare facility location. Comput Oper Res 79:223–263

Akella M, Delmelle E, Batta R, Rogerson P, Blatt A (2010) Optimal adaptive facility location using geostatistics. Geogr Anal 42:227–244

Boonmee C, Arimura M, Asada T (2017) Facility location optimization model for emergency humanitarian logistics. Int J Disaster Risk Reduct 24:485–498

Chen Z, Chen X, Li Q, Chen J (2013) The temporal hierarchy of shelters: a hierarchical location model for earthquake-shelter planning. Int J Geogr Inf Sci 27(8):1612–1630

Hakimi SL (1964) Optimum locations of switching centers and the absolute centers and medians of a graph. Oper Res 12(3):450–459

Hakimi SL (1965) Optimum distribution of switching centers in a communication network and some related graph theoretic problems. Oper Res 13(3):462–475

Hess SW, Weaver JB, Siegfeldt HJ, Whelan JN, Zitlau PA (1965) Nonpartisan political redistricting by computer. Oper Res 13(6):998–1006

Hodgson MJ (1981) The location of public facilities intermediate to the journey to work. Eur J Oper Res 6(2):199–204

Irawan CA, Salhi S (2015) Solving large p-median problems by a multistage hybrid approach using demand points aggregation and variable neighborhood search. J Glob Optim 63:537–554

Jia T, Tao H, Qin K, Wang Y, Liu C, Gao Q (2014) Selecting the optimal healthcare centers with a modified P-median model: a visual analytic perspective. Int J Health Geogr 13:42

Klastorin TD (1985) The p-median problem for cluster analysis: a comparative test using the mixture model approach. Manag Sci 31(1):84

Li R, Tong D (2017) Accessibility maximization for facility siting: an incorporation of activity space and trip chaining. Socio Econ Plan Sci 60:1–14

Mack J, Tong D (2015) Characterizing the spatial and temporal patterns of farmers' market visits. Appl Geogr 63:43–54

Marianov V (2017) Location models for emergency service applications. INFORMS Tutorials Oper Res 11:237–262

Mcguckin N, Zmud J, Nakamoto Y (2005) Trip-chaining trends in the United States: understanding travel behavior for policy making. Transp Res Rec J Transp Res Board 1917:199–204

Meskarian R, Penn ML, Williams S, Monks T (2017) A facility location model for analysis of current and future demand for sexual health services. PLoS One 12(8):e0183942. https://doi.org/10.1371/journal.pone.0183942

Mu W, Tong D (2018) A spatial-knowledge-enhanced heuristic for solving the p-median problem. Trans GIS 22:477–493

Murawski L, Church RL (2009) Improving accessibility to rural health services: the maximal covering network improvement problem. Socio Econ Plan Sci 43(2):102–110

Ndiaye M, Alfares H (2008) Modeling health care facility location for moving population groups. Comput Oper Res 35:2154–2161

Pamučar D, Vasin LJ, Atanasković P, Miličić M (2016) Planning the city logistics terminal location by applying the green p-median model and type-2 neuro-fuzzy network. Comput Intell Neurosci Article ID 6972818. https://doi.org/10.1155/2016/6972818

Park C, Sohn SY (2017) An optimization approach for the placement of bicycle sharing stations to reduce short car trips: an application to the city of Seoul. Transp Res A Policy Pract 105:154–166

Rahman S, Smith DK (2000) Use of location-allocation models in health service development planning in developing nations. Eur J Oper Res 123:437–452

ReVelle CS, Swain RW (1970) Central facilities location. Geogr Anal 2(1):30–42

ReVelle CS, Eiselt HA, Daskin MS (2008) A bibliography of some fundamental problem categories in discrete location science. Eur J Oper Res 184:817–848

Tong D, Murray AT (2017) Location analysis: developments on the horizon. In: Jackson R, Schaeffer P (eds) Regional research frontiers, vol 2. Springer International Publishing, New York, pp 193–208

Tong D, Ren F, Mack J (2012) Locating Farmers' Markets with an incorporation of spatio-temporal variation. Socio Econ Plan Sci 46(2):149–156

Widener MJ, Horner MW (2011) A hierarchical approach to modeling hurricane disaster relief goods distribution. J Transp Geogr 19(4):821–828

Widener MJ, Metcalf SS, Bar-Yam Y (2012) Developing a mobile produce distribution system for low-income urban residents in food deserts. J Urban Health 89(5):733–745

Wu C, Murray AT (2005) Optimizing public transit quality and system access: the multiple-route, maximal covering/shortest-path problem. Environ Plann B Plann Des 32:163–178

Zhang Y, Murray A, Turner BL (2017) Optimizing green space locations to reduce daytime and nighttime urban heat island effects in Phoenix, Arizona. Landsc Urban Plan 165:162–171

Chapter 5
Interpreting the Geography of Human Capital Stock Variations

Rachel S. Franklin

Abstract A wealth of research has documented the importance of human capital for economic growth and development. While much of this body of research focuses on estimating the relationship between some economic outcome and, generally, levels of educational attainment, a subsidiary corpus of research has developed that focuses on documenting and explaining the geographic variation in human capital stocks that exists. The popular press, in its turn, has also adopted human capital stocks as a proxy for urban and regional vibrancy. Little attention has been focused on what, in fact, constitutes a talent or human capital magnet and how different measures of a seemingly straightforward concept might not only generate different results but might also be capturing more than simply levels of educational attainment. This chapter uses data on educational attainment—the share of the population with at least a college degree—for US metropolitan areas in 2000 and 2010 to conceptualize what is meant by a human capital or talent magnet and to highlight a few ways in which results might be driven by definition and measure. Of particular interest are the roles of age structure, migration, and relative performance.

Keywords Human capital · Labor markets · Talent · Migration · Age structure · Shift-share analysis

5.1 Introduction

In early September 2017, Amazon released a request for proposals from North American regions and metropolitan areas to bid for its second headquarters (HQ2) location. Along with specifications regarding site and transportation requirements, the proposal request noted a preference for, "Urban or suburban locations with the

R. S. Franklin (✉)
Centre for Urban and Regional Development Studies (CURDS), Newcastle University, Newcastle upon Tyne, UK
e-mail: rachel.franklin@newcastle.ac.uk

R. S. Franklin (ed.), *Population, Place, and Spatial Interaction*, New Frontiers in Regional Science: Asian Perspectives 40,
https://doi.org/10.1007/978-981-13-9231-3_5

potential to attract and retain strong technical talent" (p. 1, Amazon 2017). The proposal later more specifically states that: "A highly educated labor pool is critical and a strong university system is required" (ibid. p. 5). On one hand, this requirement speaks directly to Amazon's eventual need for a highly-trained labor force to provide the estimated 50,000 new employees it anticipates hiring at HQ2. Certainly, a firm cannot be expected to locate in a place lacking the requisite human capital. Other language in the request, though, suggests that Amazon views existing talent (and a city's ability to attract that talent) as a synecdoche or proxy for a bundle of related characteristics that would make an area a good "cultural community fit" (ibid. p. 5) for the company.

Put another way, human capital stocks are viewed as the embodiment of a range of observable and unobservable characteristics of an area. That this presumption of the power of human capital stocks has been so widely adopted is the result of decades of research in economic geography, regional science, and economics. However, enthusiasm for the finding—that more human capital leads to better outcomes for cities and regions—may have inadvertently distracted from a more basic, but fundamental question: How do we measure human capital stocks for areas in the first place? And to what extent will our conclusions vary if we adjust the measurement? Answers to these questions are valuable to researchers, but they can also clearly drive perceptions of metropolitan areas, as well as decisions by individuals, firms, and policymakers.

There is a compelling logic to human capital research. At the individual level, increased human capital is associated with higher productivity and earnings (Becker 1964). In the aggregate, human capital is important for cities, labor markets, regions, and nations, as it is one of the building blocks of economic growth and development (Lucas 1988). This in turn leads to a natural interest in the geography of human capital (or talent, as Florida (2002) terms it): How do we explain why some places have more human capital than others, and, equally importantly, how do we understand why some places increase their human capital stocks more than others over a given period of time? If talent is unevenly distributed, it follows that this is the result of individuals "voting with their feet" and choosing to migrate or stay in place, in response to economic opportunity or desirable locations—the key element being the mobility of human capital. If this is the case, and human capital is choosing where to locate, cities and regions can be compared and ranked based on their existing (and changing) human capital stocks.

Two issues relevant to the present chapter run throughout this wide-ranging body of research. The first is encountered with both individual and urban/regional research and is the way in which human capital is measured. Although we understand the concept of human capital to be multidimensional and to include not only education, skills, and training but also unobservable qualities such as emotional intelligence, the way in which it is typically captured in data is with educational attainment—often the possession of a college degree. The second issue relates to how the educational attainment of an area is measured. This issue is particular to research that focuses on the human capital stocks of an area, whether in order to understand some outcome, such as economic or population growth; to compare the relationship between, say, city characteristics and human capital; or to compare human capital endowments across a set of locations. The straightforward approach to summarizing the educational attainment of an area is to employ raw numbers or, more often, to calculate the share of the adult population possessing at least a college degree. There are shortcomings to these statistics, however, in terms of how such numbers should be interpreted. It is this second issue that is the focus of this chapter.

The pitfalls related to characterizing educational attainment are most apparent in studies that seek to compare and rank areas—usually cities or metropolitan areas—based on their human capital stocks. Although eschewed by the bulk of the peer-reviewed literature, such rankings are common in policy reports and the popular press and are precisely the sorts of information brought to bear in discussions such as Amazon's HQ2 site selection. Here, the compelling logic to human capital research outlined above is important: comparisons of human capital stocks across cities represent outcomes of assumptions about the relationship between growth and human capital, urban desirability and human capital, or future development prospects and human capital. Following the logic, observing human capital stocks and changes to those stocks is sufficient for making pronouncements about the relative health and vitality of urban areas.

The questions are, then, what are the choices for characterizing a metropolitan area's human capital stocks, what are the shortcomings of those measurement options, and how do conclusions vary by type of measure? Importantly, this chapter argues that no one perfect measure exists but that some general guidelines do emerge. Changes in human capital stocks should not be confused with migration trends, for example. Age structure, too, is an important factor that merits attention. Finally, a distinction can be made between measures that capture the "magnet in space" aspect of human capital stocks, those that highlight change over time, and those that seek to isolate the sources of increases in human capital stocks across metropolitan areas. Using data on educational attainment and population change in US metropolitan areas for the early 2000s, this chapter compares and assesses a range of approaches to measuring human capital, showing along the way how different measures elicit different information. Some metropolitan areas perform well regardless of measure; for others, measurement seems to matter.

Seven propositions are considered. These are not intended as exhaustive but rather are reflective of a range of caveats and assumptions that underlie research on the geography of human capital in the United States and elsewhere. Following a description of related literature and the data used in this study, the following questions are evaluated:

1. What is the appropriate human capital benchmark?
2. What do rankings of human capital stocks tell us?
3. How do we frame our expectations of "expected" levels of human capital?
4. What does it mean when shares of the educated increase over time?
5. Why not measure attractiveness directly, via migration?
6. What about age structure?
7. How can we isolate increases in human capital due to area competitiveness?

5.2 Background and Data

5.2.1 Background Literature

The main reason we care about the geography of human capital—with "we" referring primarily to researchers and policymakers—is that the literature showing a positive connection between educated or skilled workers and economic growth is so extensive as to almost defy citation. It is, in fact, taken as a given in the present day that, although there may be downsides to highly educated cities, these areas tend to perform better in the long run, in terms of wages and economic and population growth (Romer 1986; Lucas 1988; Glaeser 1998; Glaeser et al.1995; Glaeser and Saiz 2004; Moretti 2004; Bauer et al. 2012).

This work, largely undertaken by economists, opens the field to what many (especially those reading this chapter) would consider the truly interesting question: if human capital matters, how do we explain its uneven distribution over space? Economists touch on this question, of course, as it relates to urban agglomeration and the desire of skilled workers to be located close to others of their kind, but also as it results from shifting geographies of production (Storper and Scott 2009; Moretti 2012). Arguably, though, the interesting considerations of the nuances of the geography of human capital come out of the spatial disciplines of geography and regional science. One of the most important of these is the distinction between human capital as it accumulates through migration or through alternative mechanisms. Although migration is the larger of the two literatures and merits separate coverage below, the other ways in which areas acquire human capital offer important food for thought. Brown et al. (2010), in their study of changes in human capital across Canadian cities, distinguish between in situ human capital production and increases due to both domestic and international migration. Any statistic showing the educational attainment of the population will reflect not only migration but also the area's ability to educate and then retain that population. What this suggests, where the present chapter is concerned, is that migration and educational attainment statistics should not be conflated. Other related research, including McHenry (2014), includes the intergenerational transfer of human capital as one factor that explains geographic persistence of human capital distribution across American commuting zones. Abel and Deitz (2012), in turn, investigate the contribution of higher education institutions (which provide in situ human capital production) to variations in human capital levels across the United States. They find that, although the presence of colleges and universities helps explain why some metropolitan areas have higher stocks than others, migration into metropolitan areas is also important.

Indeed, migration of the educated represents a sizable share of the research on human capital distribution. The likelihood of moving increases with educational attainment, and the educated tend to move longer distances than those with less than a college degree (Franklin 2003). Most important, in the context of regional development and human capital, is the fact that the benefits of migration are hypothesized to accrue not only to the individual but also to the region or city to which they move.

This has led to a great deal of research that seeks to understand the migration behavior of this group, in terms of propensity, destination choice, and benefits to areas that attract this group (e.g., Faggian et al. 2007; Whisler et al. 2008). It has also generated interest in understanding the migration behavior of those departing for university education, in order to consider the effects of this group on the destination region postgraduation (Faggian and McCann 2006; Faggian and Franklin 2014; Franklin and Faggian 2014). The young and educated are of particular interest because they can be viewed as freer to move than their older counterparts. They may also be more selective in their destination choice and have more years in the labor force ahead of them than older educated workers. See Corcoran and Faggian (2017) or Faggian et al. (2017) for very recent coverage of the subject of graduate migration. The freedom to "vote with their feet" that is attributed to skilled workers, and especially young skilled workers, translates into strong interest at the metropolitan level in knowing which areas are most attractive to this group.

With the exception of research on the migration of the young and educated, an understudied aspect of migration and human capital distribution is the role of age and age structure. Age is important because migration varies a great deal over the life course, in terms of both propensity and destination choice (Plane and Heins 2003; Plane et al. 2005). Age structure is important from a migration perspective because large cohorts of a particular age can imply large impacts on the migration system as a whole, if movement is depressed or increased for those groups (Plane and Rogerson 1991a). Age and age structure are also vitally important, however, for interpreting measures of human capital stock, as educational attainment also varies greatly by age. In general, places that are older will also be less educated, as educational attainment rates are lower for older cohorts (e.g., those over 60 years of age). Educational attainment has more or less increased over time in the United States and elsewhere, such that younger places, holding other factors constant, can be expected to be more educated than older places.

A complementary approach to understanding the geography of human capital is to hone in on the qualities of the destinations, rather than, say, the migrants. To approach the subject from this direction is generally to more explicitly ask what sorts of characteristics of places or metropolitan areas attract the educated, acknowledging that preferences may vary by demographic, stage of the life course, or age cohort (Ruth and Franklin 2014). This is done in Whisler et al. (2008) and for cores and peripheries of metropolitan areas in Walker (2017) but is perhaps most closely associated with the work of Richard Florida. Much of Florida's work has linked the "creative class" or "talent" to features of the urban landscape, such as amenities, coolness, or diversity (e.g., Florida 2002, 2014; or Florida et al. 2008). The argument put forth has been that the unevenness of the distribution of human capital can be at least partially explained by variations in the sorts of characteristics listed earlier. Educated individuals, who have a choice in where to work and live, will opt for those metropolitan areas that most appeal to their preferences. For the purposes of this chapter, what matters is the implicit assumption that human capital—talent—can be attracted from one place to another. Following the compelling logic of human capital research, skilled workers are good for economies, they vote with their feet, and they prefer certain types of destinations. Translated into lay language and policy guidance, this suggests that metropolitan areas can be compared and judged based on the size of their educated population. Other related research discussed earlier, though, also makes clear that although the temptation exists to ascribe changes in human capital stocks to migration (and therefore attractiveness of a city), there are other factors—including age structure and in situ production of human capital—that may explain both variations in measures across places as well as observed changes over time. Examples of these hazards are outlined later in Sect. 5.3.

5.2.2 Data

As in previous research, this chapter uses educational attainment as its measure of human capital.[1] Expanding on previous research, which has tended to focus on the educational attainment of the population as a whole, or least some large segment thereof (e.g., McHenry 2014 uses the percent of the population 24–64, whereas Florida 2002 appears to employ the share of the total population with a degree), this chapter looks at educational attainment by age cohort, as well as for the population 25 and up. County-level data on education attainment and total population come from the Census 2000 long form and from the American Community Survey (ACS) 2009–2013 estimates. For the sake of simplicity, the 2009–2013 estimates are henceforth referred to as 2009. All data files were retrieved from the University of Minnesota's IPUMS National Historical Geographic Information System (Manson et al. 2017). County data for both time periods are aggregated to 2013 core-based statistical areas, or CBSAs—the formal term for American metropolitan areas. All analysis and tables below refer to results for metropolitan areas, of which there were 381.

5.3 What Makes a Talent Magnet?

The main argument of this chapter is that, for various reasons, adopting the share of the population that is college educated as the primary measure of human capital can lead to erroneous conclusions about the distribution of human capital and the level of attractiveness of metropolitan areas to the educated. The following section addresses the shortcomings of this default statistic, provides alternative measures, and shows how conclusions may vary depending on the way in which an area's human capital stocks are assessed.

[1]In fact, this chapter uses the terms "human capital," "talent," "educated," and "skilled" interchangeably.

Table 5.1 United States: percent of the population possessing at least a college degree

Year	Geography	Total population		Population, by age cohort				
		18+	25+	18–24	25–34	35–44	45–64	65 +
2000	United States	22.26	24.4	7.81	27.54	25.88	26.39	15.39
	US metro areas	23.91	26.22	8.48	29.43	27.8	28.22	16.56
2009	United States	26.3	28.84	9.43	31.94	32.27	28.89	22.3
	US metro areas	28.06	30.8	10.12	33.94	34.3	30.82	23.83

5.3.1 What Is the Appropriate Benchmark?

As a point of departure, it is helpful to establish a few benchmarks for human capital stocks in the United States. Without these, it is difficult to assess whether any place in the country has "high" or "low" amounts of human capital. Table 5.1 shows the share of the US population with at least a college degree for two time periods, 2000 and 2009–2013, for the nation as a whole, as well as for metropolitan parts of the country. Two unsurprising conclusions may immediately be drawn from this table. First, human capital stocks tend to increase over time. In 2000, almost a quarter of the population 25 and up possessed a college degree. By 2009, this share had increased to almost 29%. In fact, over the longer run, this is the most educated US population that has ever been. In 1950, only 6.2% of the population had a college degree, and by 1990, this number stood at just over 20% (US Census Bureau 2006). The stark increase over time reflects increasing educational attainment—larger shares of the population were more likely to attend university—but also the inevitable dying out of older and less educated cohorts, who are then replaced by the more educated. The second conclusion is that metropolitan areas are more educated than the population as a whole, and this holds across individual age cohorts, as well.

National-level statistics are also a helpful starting place for considering the importance of age. Table 5.1 also shows educational attainment by age cohort. In both time periods, working-age cohorts are more educated than the metropolitan total, indicating the effect of the 65+ cohort's lower educational attainment on the global measures. Within each cohort, educated shares increased over the time period, reflecting cohort-level changes in educational attainment. That is, the increase for the 65 and up cohort from 16.56% to 23.83% between 2000 and 2009 is not of course an indication of more seniors deciding to attend college; rather, it shows the importance of cohort succession over time. Those aging into this age cohort are more educated than those leaving it. How to explain the sharp increase in educational attainment of the 35–44 cohort in 2009? This is the 25–34 cohort of 2000—meaning that their education attainment climbed from 29.43% to 34.3% over the decade. Some of this is likely due to the 5-year ACS estimates used for the later time period. Otherwise, the increase can be attributed not only to the members of the cohort completing their college education but also likely to international immigration. The main point is that, given prevailing increases in educational attainment over time, the educated shares of population, both at the aggregate level and for individual cohorts, *should* increase between time periods.

Table 5.2 Percent of group that has college degree, largest metropolitan areas, 2009–2013

Metropolitan area	Total population 2009–2013	Age cohort				
		25+	25–34	35–44	45–64	65+
New York-Newark-Jersey City, NY-NJ-PA	19,716,880	36.47	44.88	40.82	35.40	25.14
Los Angeles-Long Beach-Anaheim, CA	12,945,252	31.36	33.68	32.52	31.27	27.10
Chicago-Naperville-Elgin, IL-IN-WI	9,488,493	34.58	41.19	38.66	33.83	23.37
Dallas-Fort Worth-Arlington, TX	6,575,833	31.72	31.25	33.69	32.83	26.44
Houston-The Woodlands-Sugar Land, TX	6,063,540	29.43	29.19	30.43	30.46	25.45
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	5,992,766	33.57	39.97	38.81	33.18	23.14
Washington-Arlington-Alexandria, DC-VA-MD-WV	5,759,330	47.85	51.49	51.19	47.27	39.18
Miami-Fort Lauderdale-West Palm Beach, FL	5,673,185	28.91	29.25	31.69	30.20	24.13
Atlanta-Sandy Springs-Roswell, GA	5,379,176	34.93	35.86	39.25	35.34	25.41
Boston-Cambridge-Newton, MA-NH	4,604,278	43.40	54.37	49.42	42.14	28.54
San Francisco-Oakland-Hayward, CA	4,402,729	44.53	49.57	50.49	42.80	35.40
Detroit-Warren-Dearborn, MI	4,295,700	28.08	31.15	33.34	28.23	19.88
Riverside-San Bernardino-Ontario, CA	4,285,443	19.65	18.36	19.73	20.18	20.06
Phoenix-Mesa-Scottsdale, AZ	4,268,289	28.72	27.48	30.58	29.62	26.49
Seattle-Tacoma-Bellevue, WA	3,504,628	38.00	40.01	41.78	37.41	31.71

The established approach to judging an area's human capital is to assume that all places are competing for a fixed number of skilled or educated individuals. While this is true in many ways, it leads immediately to the ranking of areas, whereby more is always better. An alternative view is to judge a metropolitan area's human capital stocks in relation to some benchmark—the comparable figure for all metropolitan areas, for example. Comparing the national metropolitan figures for 2009 to data for the largest US metropolitan areas reinforces some common views about talent magnets in the United States (Table 5.2). Washington, DC, San Francisco, and Boston, for example, all emerge as well above the national metropolitan benchmark. Areas such as Houston and Phoenix by this definition, however, would not qualify as magnets, looking at the 25 and up category. Within particular age cohorts, some areas, such as Philadelphia, appear to be especially attractive to younger age cohorts. Atlanta, though, is only slightly above the benchmark for the 25–34 cohort, suggesting that its reputation as a quick-growing, dynamic city does not necessarily hold for those just out of college.

Table 5.3 Most educated metropolitan areas, 2009

Metropolitan area	Percent educated, 25+	Metropolitan area	Percent educated, 25–34
Boulder, CO	58.30	Ames, IA	84.35
Ann Arbor, MI	51.26	Manhattan, KS	82.17
Corvallis, OR	49.96	Ithaca, NY	78.45
Lawrence, KS	49.56	Lawrence, KS	77.85
Ithaca, NY	49.29	Iowa City, IA	75.95
Washington-Arlington-Alexandria, DC-VA-MD-WV	47.85	Columbia, MO	72.19
Ames, IA	47.72	Champaign-Urbana, IL	68.60
Columbia, MO	47.28	Gainesville, FL	67.06
Iowa City, IA	46.23	Ann Arbor, MI	64.05
San Jose-Sunnyvale-Santa Clara, CA	45.70	Bloomington, IN	63.43
Bridgeport-Stamford-Norwalk, CT	44.78	State College, PA	62.20
San Francisco-Oakland-Hayward, CA	44.53	Corvallis, OR	60.34
Durham-Chapel Hill, NC	44.05	College Station-Bryan, TX	57.89
Fort Collins, CO	43.79	Missoula, MT	57.28
Boston-Cambridge-Newton, MA-NH	43.40	Madison, WI	56.83

5.3.2 *What Do Rankings of Human Capital Stocks Tell Us?*

When all metropolitan areas are ranked by the percent educated in order to identify talent magnets, the magnets that emerge are, on the whole, a particular group of areas. Again, San Francisco, Washington, DC, and Boston emerge as magnets. Their peers, however, are recognizable as being mostly college towns. Among the population 25 and up, Boulder, Colorado, tops the list with almost 60% of its adult population possessing a college degree. For the 25–34 cohort, Ames, Iowa, home of Iowa State University, is at the top of the rankings: over 80% of its young adults were college educated in 2009. In fact, for the young adult cohort, all of the top 15 educated metropolitan areas are known primarily as the homes of large state universities in smaller cities and metropolitan areas. The relative number of people living in these areas, compared to the larger metropolitan areas, is small. Even if these places qualify as talent magnets given usual definitions (i.e., the share of the population that is educated), they are small magnets whose attractive field cannot be imagined to compare with those of bigger areas.

Table 5.3's results, then, suggest that, in identifying magnets, primary interest lies not only in the share of the population that is educated but also in whether this number is higher than otherwise expected—and also how large the magnet is. Where college towns are concerned, they are not only smaller, but it is also expected that most of the inhabitants will be college educated, whether graduate students, faculty, support staff, or even the sort of older individual who prefers a college town retirement destination.

5.3.3 Which Areas Have More of the Educated Than We Might Expect?

At any given point in time, some measurable quantity of college-educated individuals is spread across US metropolitan areas. One type of magnet is certainly those areas with the largest shares of the educated overall. We could think of these metro areas as the quintessential "magnet" in that these are the locations that act as poles or epicenters of human capital. By this standard, the New York metropolitan area was the clear forerunner in both 2000 and 2009, with almost 10% of all metropolitan human capital. Los Angeles, Chicago, and Washington, DC, follow New York, with about 3–5% of all the metropolitan educated 25 and up. The results are the same if age is restricted to only the metropolitan educated ages 25–34. In both periods, fifth place for the 25+ population is held by San Francisco, whereas Boston is fifth for the 25–34 age cohort. These numbers indicate something about the lumpiness of human capital distribution across all metropolitan areas (in fact, the lowest-ranking areas have only hundredths of a percent of all the metropolitan human capital and so easily qualify as not magnets), but since these are among the largest urban areas in the country, it is not surprising that they also possess large quantities of the educated.

In order to identify those metropolitan areas with more human capital than expected, we can turn to one of the workhorse statistics of geography and regional science, the location quotient. Essentially a ratio of ratios, the location quotient in this case can highlight those areas where the metropolitan area's share of the educated population is greater than its share of the total population in that age cohort, so:

$$LQ_{HC25+}\frac{CBSA\ Educated\ Population\ 25+/National\ Educated\ Population\ 25+}{CBSA\ Population\ 25+/National\ Population\ 25+}$$

Areas with location quotients for human capital over one are those that have more of the educated than their share of the total population in that category might indicate. Quotients under one suggest underrepresentation of the educated. Table 5.4 presents human capital location quotients for 2009 for the selection of metropolitan areas with the largest shares of their population in each age category, either 25 and up or 25 to 34. By this standard of talent magnet assessment, many of the same metropolitan areas rise to the top: Washington, DC, Boston, and San Francisco are all very well endowed with the educated, relative to the total population in each age group living in those areas. New York and Chicago also perform well. Among the 25 and up group, Minneapolis emerges as a possible magnet, while among the 25–34 category, Denver and Baltimore stand out. Other areas, such as Los Angeles, St. Louis, or Dallas, have about the amount of human capital that could be expected (i.e., location quotients around one). Atlanta falls into this category for the 25–34 group, with a location quotient of 1.06. Finally, there are the areas such as Phoenix, Detroit, or Tampa, which are less educated than expected, given this measure.

Table 5.4 Human capital location quotients, 2009[a]

Metropolitan area	Population 25+	Metropolitan area	Population 25–34
New York-Newark-Jersey City, NY-NJ-PA	1.18	New York-Newark-Jersey City, NY-NJ-PA	1.32
Los Angeles-Long Beach-Anaheim, CA	1.02	Los Angeles-Long Beach-Anaheim, CA	0.99
Chicago-Naperville-Elgin, IL-IN-WI	1.12	Chicago-Naperville-Elgin, IL-IN-WI	1.21
Dallas-Fort Worth-Arlington, TX	1.03	Dallas-Fort Worth-Arlington, TX	0.92
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	1.09	Houston-The Woodlands-Sugar Land, TX	0.86
Miami-Fort Lauderdale-West Palm Beach, FL	0.94	Washington-Arlington-Alexandria, DC-VA-MD-WV	1.52
Washington-Arlington-Alexandria, DC-VA-MD-WV	1.55	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	1.18
Houston-The Woodlands-Sugar Land, TX	0.96	Atlanta-Sandy Springs-Roswell, GA	1.06
Atlanta-Sandy Springs-Roswell, GA	1.13	Miami-Fort Lauderdale-West Palm Beach, FL	0.86
Boston-Cambridge-Newton, MA-NH	1.41	San Francisco-Oakland-Hayward, CA	1.46
San Francisco-Oakland-Hayward, CA	1.45	Boston-Cambridge-Newton, MA-NH	1.60
Detroit-Warren-Dearborn, MI	0.91	Phoenix-Mesa-Scottsdale, AZ	0.81
Phoenix-Mesa-Scottsdale, AZ	0.93	Riverside-San Bernardino-Ontario, CA	0.54
Riverside-San Bernardino-Ontario, CA	0.64	Seattle-Tacoma-Bellevue, WA	1.18
Seattle-Tacoma-Bellevue, WA	1.23	Detroit-Warren-Dearborn, MI	0.92
Minneapolis-St. Paul-Bloomington, MN-WI	1.25	Minneapolis-St. Paul-Bloomington, MN-WI	1.29
San Diego-Carlsbad, CA	1.12	San Diego-Carlsbad, CA	1.04
Tampa-St. Petersburg-Clearwater, FL	0.86	Denver-Aurora-Lakewood, CO	1.20
St. Louis, MO-IL	1.00	Baltimore-Columbia-Towson, MD	1.21
Baltimore-Columbia-Towson, MD	1.17	St. Louis, MO-IL	1.09

[a]Sorted by metropolitan area share of population in that age category

5.3.4 *What Does It Mean When Shares of the Educated Increase Over Time?*

One way of thinking about the above measures, location quotients or the educated share of population, is as magnets in space—for a given point in time, they assess the existing geography of human capital. Another perspective is gained by looking at changes over time: which areas increase their human capital stocks the most? This approach is not incompatible with the measures above; location quotients, for example, could easily be compared over time. The difficulty is the one outlined above, that populations *should* become more educated over time. Thus, the challenge is to identify those areas that increase their human capital stocks more than might be expected, given changes in age structure and educational attainment. One approach could be to identify areas that increased at a higher rate than the nation as a whole or all metropolitan areas. Alternative options are to rank areas that experienced the highest increases in their educated population over time or to compare the change in educated to the change in the total population, similar to the location quotient approach. For illustrative purposes, these statistics are calculated for metropolitan areas of over one million in 2009 and then ranked by the percent change in the educated for the entire adult population (i.e., 25 and up).

Unlike results presented above, the figures shown in Table 5.5 privilege those areas with the greatest improvements in their human capital stocks—in many cases, their educated populations pale in comparison to the talent magnets already identified above. Las Vegas, for example, increased its human capital stocks by over 85% between 2000 and 2009; however, only just over 22% of the population 25 and up had a college degree in 2009. Other areas that are occasionally touted as talent magnets do appear on the list, however: Columbus, Ohio; Austin, Texas; and Charlotte, North Carolina, for example. Atlanta, which performs solidly according to other measures, also appears on this list.

The young and educated are often the focus of studies on the geography of talent. Table 5.5 shows how human capital for this age group changed in the large metropolitan areas that are increasing their overall human capital stocks the most. In every case, increases in the educated outstripped increases in population growth in the 25–34 age cohort. This mainly reflects the higher levels of educational attainment of those moving into this age cohort, relative to those who moved on to the 35–44 cohort. For some of these areas, increases may also reflect larger numbers of the educated remaining in place, rather than moving elsewhere. For some places, Atlanta and Austin in particular, increases in both statistics are similar, suggesting that, as these cohorts grew, they were mainly college-educated individuals. Interestingly, Atlanta only increased its stock of 25–34-year-olds by about 3% between 2000 and 2009. This is striking for a metropolitan area that grew by about 25% during this time period.

Table 5.5 Percent change in college educated, large metropolitan areas, 2000–2009[a]

Metropolitan area	Total population 2009	Population 25+		25–34 cohort	
		Numerical change, educated	Percent change, educated	Percent change, educated	Percent change, population
Las Vegas-Henderson-Paradise, NV	1,976,925	133,190	85.33	65.67	32.19
Riverside-San Bernardino-Ontario, CA	4,285,443	200,706	64.28	81.39	32.21
Raleigh, NC	1,162,689	124,553	64.22	25.11	19.63
Austin-Round Rock, TX	1,782,032	180,847	64.17	40.01	33.68
Charlotte-Concord-Gastonia, NC-SC	2,261,321	175,031	62.30	27.80	10.74
San Antonio-New Braunfels, TX	2,192,724	129,883	55.86	54.00	27.11
Jacksonville, FL	1,363,610	89,282	54.21	45.41	14.90
Nashville-Davidson-Murfreesboro-Franklin, TN	1,702,603	120,663	53.62	34.35	16.72
Phoenix-Mesa-Scottsdale, AZ	4,268,289	274,280	53.25	32.26	18.51
Orlando-Kissimmee-Sanford, FL	2,183,363	136,547	50.81	35.47	25.61
Houston-The Woodlands-Sugar Land, TX	6,063,540	359,388	47.27	43.72	24.05
Portland-Vancouver-Hillsboro, OR-WA	2,260,591	168,155	46.36	29.35	13.38
Dallas-Fort Worth-Arlington, TX	6,575,833	397,614	43.20	20.56	11.40
Salt Lake City, UT	1,107,434	61,384	42.90	47.52	24.84
Sacramento-Roseville-Arden-Arcade, CA	2,174,401	128,582	42.53	33.52	20.12
Tampa-St. Petersburg-Clearwater, FL	2,819,241	156,209	42.51	35.32	14.31
Atlanta-Sandy Springs-Roswell, GA	5,379,176	357,676	41.84	5.38	3.02
Denver-Aurora-Lakewood, CO	2,601,465	199,555	41.39	26.95	13.37
Columbus, OH	1,926,242	120,013	40.64	23.67	7.30
Louisville/Jefferson County, KY-IN	1,244,880	62,152	39.00	29.07	6.10

[a]Sample is metropolitan areas of one million or more, sorted by percent change in college educated, 25 and up, 2000–2009

5.3.5 Why Not Measure Attractiveness Directly, via Migration?

Part of the appeal of identifying temporal talent magnets—those areas that gain the most human capital over a given period of time—is the hope that dynamic measures are capturing something about the relative appeal of various metropolitan areas. Change over time provides a good indication of relative desirability, but also possibly a region's commitment to education and training, as well as provision of a high quality of life (which prompts human capital to remain in place). In practice, what often occurs is that all increases in human capital over time are ascribed to migration. Cities and metropolitan areas with the largest increases in the shares of their population with a degree are judged to be the most attractive or desirable—talent magnets, in short. The issue with this type of measurement is that, as noted, shares *should* increase over time, and increases can result from demographic changes other than migration.

Ideally, stocks and flows of human capital would be disaggregated for metropolitan areas. A possible gold standard for identifying talent magnets would be in- and out-migration data for all metropolitan areas on a regular basis (similar to Burd 2013 or Walker 2017). Such an approach would highlight those areas with the highest net inflows of the educated and would easily pinpoint those metropolitan areas that attract the most migrants. Disaggregated flows that permitted researchers to discover where the educated move from and to would also help clarify how different types of areas grow their human capital. Some areas, for example, might attract large numbers of the educated from other metropolitan areas, while other places might primarily attract from their metropolitan hinterland or internationally. Use of migration data is not the standard approach, but it is included here as it is one of the measures of talent magnets that most closely resembles the concept researchers are often trying to capture.

5.3.6 What About Age Structure?

Up to this point, much has been made of the importance of age structure for the measurement of human capital stocks. In this section, for the sake of clarity, a simple model is constructed that shows how age structure can directly impact the usual calculations of educational attainment for an area. Here, two hypothetical cities are compared, both with the same total population and both subject to the same age-specific rates of educational attainment and mortality. Population distributions for each city are fictional but can be thought of as a younger city (City A) such as Provo, Utah, and an older city (City B), such as Toledo, Ohio (Table 5.6).

In Time 1, each age cohort's educational attainment is allocated according to rates from Census 2000 (see Table 5.8). Because younger cohorts are more educated than older cohorts, the whole City A is more educated than City B—simply because City

Table 5.6 Hypothetical example of age structure impact

	City A				City B			
	Time 1		Time 2		Time 1		Time 2	
Age cohort	Total	Educated	Total	Educated	Total	Educated	Total	Educated
25–29 years	10,000	2715	10,000	3000	5000	1358	5000	1500
30–34 years	10,000	2791	9952	2702	5000	1396	4976	1351
35–39 years	7500	1943	9945	2776	7500	1943	4973	1388
40–44 years	7500	1940	7448	1929	7500	1940	7448	1929
45–49 years	5000	1423	7425	1921	5000	1423	7425	1921
50–54 years	5000	1456	4920	1400	5000	1456	4920	1400
55–59 years	2500	616	4878	1420	7500	1848	4878	1420
60–64 years	2500	508	2412	594	7500	1524	7237	1783
Total working age	50,000	13,391	56,981	15,742	50,000	12,886	46,857	12,692
Percent educated		26.78		27.63		25.77		27.09
Change in number educated				17.56				−1.51

A has more individuals in younger age cohorts, while City B has larger older age cohorts. The result is that, in Time 1, almost 27% of City A's population is educated, compared to almost 26% of City B's adult population.

Between Time 1 and Time 2, two things happen, holding other changes constant. First, some members of each age cohort will die (see Table 5.9 for cohort-level mortality risk, from Arias 2014). Here we assume that the educated and uneducated are equally likely to die; that is a strong assumption. Second, surviving individuals graduate from one age cohort to the next.[2] In both cities, the share of the population that is educated increases during the time period. In fact, City B has a higher increase in the percent of the population educated, because so many of its uneducated older individuals age out of their working years. In terms of changes to the raw numbers of educated in each city, however, City B has fewer educated in Time 2 than in Time 1 and experiences a decline. City A, though, experiences robust growth in its numbers of educated.

In reality, of course, educational attainment varies across metropolitan areas, and the educated migrate. This example highlights the contribution of age structure alone to educational attainment measures. In their way, these statistics are as vulnerable to age structure bias as other demographic measures used for fertility or mortality. At a minimum, the results suggest that, as is often done for mortality and fertility, age-specific educational attainment rates are an improvement over measures for the entire adult population.

5.3.7 How Can We Isolate Increases in Human Capital Due to Area Competitiveness?

A recurring issue in the measures discussed above, and one that is implicit in other research and coverage of the subject, is how to identify those areas that stand out as attracting or growing their human capital more than their peers. As discussed above, levels of human capital have increased over time across the United States as a whole. In addition, clearly age structure and age-specific education rates are important. How is it possible, then, to sift out the increases that are due to a metropolitan area's attractiveness?

One solution, once popular in economic geography and now perhaps more common in population geography, is to employ shift-share analysis as a descriptive tool for the evaluation of changes occurring in employment, fertility, population composition, or migration (e.g., Plane 1987; Ishikawa 1992; Franklin and Plane 2004; Franklin 2014). A variety of shift-share tools have been developed, but the

[2]For the sake of argument, the size of the incoming 25–29-year-old cohort is assumed to be the same size as the preceding cohort. Educational attainment rates for this young cohort are assumed to have increased marginally and no one who did not already have a degree acquires one as they age. In addition, no migration takes place.

basic approach is to decompose change in some quantity, whether jobs, people, or births. In the case of human capital, the interest lies with metropolitan area-level changes in the numbers of college educated over a given time period, here 2000–2009. Basic shift-share estimates three components of change. The first component, termed the National Effect, captures metropolitan area changes in human capital stocks that can be attributed to national-level forces. The Cohort Mix Effect estimates the share of total change that is due to the distribution of the educated across age cohorts. Those places that "specialize" in age cohorts that are growing robustly at the national level will benefit from their advantageous mix of age-specific educated. Finally, the Competitive Mix Effect captures growth in the educated that is attributable to the metropolitan area being competitive within a particular age cohort—that is, its own cohort educated growth rate outstripped the cohort growth rate at the national level. It is this Competitive Mix Effect which is of special interest here. Each shift-share component for area j sums to the total cohort (x) change in educated over the study period, like so:

$$\begin{aligned}&\Delta\, Educated_x^j\\&= National\ Effect_x^j + Cohort\ Mix\ Effect_x^j + Competitive\ Mix\ Effect_x^j\end{aligned}$$

Individual effects are calculated thus:

$$\begin{aligned}National\ Effect_x^j &= Educated_{x,T_1}^j \times NG\\Cohort\ Mix\ Effect_x^j &= Educated_{x,T_1}^j \times (NC - NG)\\Competitive\ Mix\ Effect_x^j &= Educated_{x,T_1}^j \times (LC - NC)\end{aligned}$$

where:

NG = National growth rate in educated, 2000–2009

NC = National cohort-level growth rate in educated, 2000–2009

LC = Local cohort-level growth rate in educated, 2000–2009

One disadvantage of shift-share analysis is that it is typically applied to one or perhaps a few areas. The analysis produces a number of results that are inconvenient for comparing large number of places. As a demonstration of the utility of the technique, Table 5.7 shows shift-share results for four areas: Atlanta, which, as shown, grew quickly during this period but may not qualify as a talent magnet; Austin, which not only experienced population growth but is also occasionally referred to as a magnet; Cleveland, which is among a set of shrinking US cities; and Columbus, which, like Austin, is sometimes proposed as a dark horse human capital magnet.

The rightmost column of Table 5.7 disaggregates the cohort contribution to total change in numbers of educated for each metropolitan area. These numbers show that, for example, although Atlanta increased its educated by over 350,000 during this period, the bulk of the growth—almost two thirds—came from the 45–64 age cohort; only a tiny fraction came from the 25–34 group. In contrast, for both Austin and Columbus, increases were more evenly spread across all age cohorts. Cleveland, though, actually saw losses in the number of educated in its younger age cohorts; all growth was due to increases in the older educated.

Table 5.7 Shift-share results, selected metropolitan areas, 2000–2009

Age cohort	Number educated, 2000	National share	Cohort mix	Competitive mix	Actual 2000–2009 change
Atlanta, Georgia					
25–34	260,510	88,592	−30,670	−43,911	14,012
35–44	258,312	87,845	−59,377	39,237	67,705
45–64	279,164	94,936	32,456	76,008	203,401
65+	56,977	19,376	22,433	30,748	72,558
Total 25+		290,750	−35,157	102,083	357,676
Austin, Texas					
25–34	88,732	30,175	−10,446	15,775	35,504
35–44	80,007	27,208	−18,391	30,152	38,969
45–64	89,971	30,597	10,460	34,256	75,313
65+	23,112	7860	9100	14,101	31,061
Total 25+		95,840	−9277	94,284	180,847
Cleveland, Ohio					
25–34	86,316	29,354	−10,162	−20,990	−1798
35–44	91,536	31,129	−21,041	−13,171	−3083
45–64	122,617	41,699	14,256	−10,686	45,269
65+	42,634	14,499	16,786	−11,141	20,144
Total 25+		116,680	−161	−55,987	60,532
Columbus, Ohio					
25–34	90,033	30,618	−10,600	1294	21,312
35–44	81,993	27,884	−18,847	9587	18,623
45–64	95,948	32,629	11,155	16,768	60,553
65+	27,351	9301	10,769	−545	19,525
Total 25+		100,432	−7523	27,104	120,013

Turning to the actual shift-share results, most growth in the educated for Atlanta and Columbus comes from the National Effect. That is, increases in the educated in these areas are not so much attributable to their being magnets as to their benefitting from increases that were affecting the country as a whole. For its part, Atlanta appears to be competitive in all age cohorts but 25–34, and was especially strong in the 65+ group (i.e., its Competitive Mix Effect far outweighs the contribution from the National Effect). Columbus, though, was only marginally competitive, with the exception of the 45–64 age group. Austin's growth is about equally the result of the Competitive Mix Effect and the National Effect. Not surprisingly, the entirety of Cleveland's growth in human capital comes from the National Effect—it is not competitive in any age cohort.

The Cohort Mix Effect captures the interplay between a metropolitan area's mix of educated across age cohorts—does the area, for example, have more older workers than fresh university graduates?—and the extent to which cohort-specific growth at the national level was greater or less than overall national growth. For example, at the national scale, the number of educated increased between 2000 and 2009 by 34%. The increase for 25–34-year-olds was only 22%, while for those 65 and up, it was 73%. So, for this time period, the Cohort Mix Effect will always be negative for the younger cohort and positive for the older cohort; what matters is how specialized the metropolitan area was in either of those groups. Those areas without many older educated did not reap as much of a benefit from the Cohort Mix Effect as those with many individuals in this group. As Table 5.7 shows, Austin gained over 30,000 educated in the 65+ category, but less than a third was attributable to the Cohort Mix Effect. In contrast, Columbus gained almost 20,000 in this group, over half of which was due to the Cohort Mix. In fact, all of Columbus's gain in the oldest educated was attributable to national or cohort-level factors; its Competitive Mix component was negative.

The main utility of shift-share analysis lies with its descriptive power, its flexibility, and its straightforward data requirements. This approach is no substitute for more fundamental measures of area-level human capital stocks. However, it is helpful for emphasizing the multiple pathways for human capital to increase in an area—for most areas, growth will occur if the benchmark region (i.e., the nation) is also growing. What makes a metropolitan area remarkable, then, is the extent to which it increases its human capital stocks over and above what might be expected, which the Competitive Mix captures.

5.4 Conclusion

There is an intense desire on the part of researchers and policymakers and, likely, Amazon executives, to know which locations are most attractive to skilled workers (i.e., the educated or talented). By its deadline, Amazon's request for HQ2 proposals had generated a phenomenal number of applications: over 200, from cities and regions across North America. For each of these interested parties, where labor is concerned, it is important to know not only how skills are distributed but also why such a spatial distribution is observed. Labor market and human capital researchers will often focus on the push and pull factors that explain this geography: amenities and other area characteristics, jobs, or in situ production of skills. The point of this chapter has been to highlight how the geography of human capital, particularly the identification of so-called magnets or hubs of the educated, depends partly on measurement and partly on definition.

The tricky aspects of human capital measurement are largely demographic. In most developed countries, younger cohorts are more educated than their elders. And, since everyone gets older 1 year at a time (a fact notably highlighted by Plane and Rogerson 1991b), the older and less educated eventually die and are replaced by the more educated. This tells us that, when we observe increases in human capital stocks over time, some portion of this share is exactly what should be expected. Comparing numbers or shares of the educated over time is also problematic when employed as a shortcut for estimating migration flows. As previous research as well as this present chapter makes clear, it is inaccurate to ascribe all increases to net in-migration. In that case, researchers and policymakers are much better off using actual migration data.

There are definitional issues related to describing the geography of human capital stocks, as well. This chapter has used the term magnet for those metropolitan areas with abundances of human capital, or college educated. But what is meant by magnet? As this chapter shows, there is a conceptual difference between those areas with the largest shares of the educated, those with more educated than expected (which in turn depends on some benchmark to provide the expectation), those with the sharpest increases in human capital stocks, and those performing relatively better than their peers. As it happens, by most measures, known epicenters of the educated—Boston, Washington, DC, and San Francisco—emerge at the top of the pile. Some of the measures discussed earlier, though, may be especially useful for identifying lagging regions or metropolitan areas. Atlanta and Columbus, Ohio, for example, may have a more complicated relationship with attracting human capital than the usual measures will indicate.

This work is intended as a conversation opener. Indeed, surely more questions have been asked than answered. The main contribution of this chapter is the argument that it matters how human capital is measured. This may lead to use of a broader range of measures in future research or at least more discussion of how to interpret analyses. It also shows the potential that exists for future research on the topic. Understanding the geography of talent, after all, is not only an intellectual exercise but also one that underpins assumptions about urban vitality and helps to drive policymaking.

Appendix

Table 5.8 Educational attainment in the United States, 2000

Cohort	In 2000	5 years later
25–29 years	0.2715	0.3*
30–34 years	0.2791	0.2715
35–39 years	0.259	0.2791
40–44 years	0.2587	0.259
45–49 years	0.2845	0.2587
50–54 years	0.2911	0.2845
55–59 years	0.2464	0.2911
60–64 years	0.2032	0.2464

*Estimated for purposes of illustration

Data Source: US Census Bureau 2006. “Census 2000 PHC-T-41. A Half-Century of Learning: Historical Statistics on Educational Attainment in the United States, 1940 to 2000”

Table 5.9 Cohort-level mortality risk, 2010

Age cohort	Probability of dying before reaching next age cohort
25–29 years	0.004791
30–34 years	0.005497
35–39 years	0.006913
40–44 years	0.009979
45–49 years	0.016044
50–54 years	0.024343
55–59 years	0.035106
60–64 years	0.049847

Data Source: Arias 2014. National Center for Health Statistics (NCHS), Table VI

References

Abel JR, Deitz R (2012) Do colleges and universities increase their region's human capital? J Econ Geogr 12:667–691. https://doi.org/10.1093/jeg/lbr020

Amazon HQ2 RFP (2017) Last accessed 7 Dec 2017 at: https://images-na.ssl-images-amazon.com/images/G/01/Anything/test/images/usa/RFP_3._V516043504_.pdf

Arias E (2014) United States life tables, 2010. Natl Vital Stat Rep 63(7)

Bauer PW, Schweitzer ME, Shane SA (2012) Knowledge matters: the long-run determinants of State income growth. J Reg Sci 52:240–255. https://doi.org/10.1111/j.1467-9787.2011.00729.x

Becker GS (1964) Human capital theory. Columbia University Press, New York

Brown WM, Newbold KB, Beckstead D (2010) Growth and change in human capital across the Canadian Urban hierarchy, 1996–2001. Urban Stud 47:1571–1586. https://doi.org/10.1177/0042098009353625

Burd, CA (2013) Metropolitan migration flows of the creative class by occupation using 3-year 2006–2008 and 2009–2011 American Community Survey Data. US Census Bureau working paper number 2013-11. Last accessed 8 Dec 2017 at: https://www.census.gov/content/dam/Census/library/working-papers/2013/demo/SEHSD-WP2013-11.pdf

Corcoran J, Faggian A (2017) Graduate migration and regional development: an international perspective. Edward Elgar Publishing, Cheltenham

Faggian A, Franklin RS (2014) Human capital redistribution in the USA: The migration of the college-bound. Spat Econ Anal 9:376–395. https://doi.org/10.1080/17421772.2014.961536

Faggian A, McCann P (2006) Human capital flows and regional knowledge assets: a simultaneous equation approach. Oxf Econ Pap 58:475–500. https://doi.org/10.1093/oep/gpl010

Faggian A, McCann P, Sheppard S (2007) Human capital, higher education and graduate migration: an analysis of Scottish and Welsh students. Urban Stud 44:2511–2528. https://doi.org/10.1080/00420980701667177

Faggian A, Corcoran J, Rowe F (2017) Special issue on youth and graduate migration. Ann Reg Sci 59:571–575. https://doi.org/10.1007/s00168-017-0845-2

Florida R (2002) The economic Geography of talent. Ann Assoc Am Geogr 92:743–755. https://doi.org/10.1111/1467-8306.00314

Florida R (2014) The rise of the creative class—revisited: revised and expanded. Basic Books (AZ), New York

Florida R, Mellander C, Stolarick K (2008) Inside the black box of regional development—human capital, the creative class and tolerance. J Econ Geogr 8:615–649. https://doi.org/10.1093/jeg/lbn023

Franklin RS (2003) Migration of the young, single, and college educated: 1995 to 2000 US Census Bureau, Census
Franklin RS (2014) An examination of the Geography of population composition and change in the United States, 2000–2010: insights from Geographical indices and a shift–share analysis. Popul Space Place 20:18–36. https://doi.org/10.1002/psp.1744
Franklin RS, Faggian A (2014) College student migration in New England: who comes, who goes, and why we might care. Northeast Geogr 6:45–60
Franklin R, Plane D (2004) A shift-share method for the analysis of regional fertility change: an application to the decline in childbearing in Italy, 1952-1991. Geogr Anal 36:1–21
Glaeser EL (1998) Are cities dying? J Econ Perspect 12(2):139–160
Glaeser EL, Saiz A (2004) The rise of the skilled city. Brookings-Wharton papers on urban affairs 2004:47–105
Glaeser EL, Scheinkman J, Shleifer A (1995) Economic growth in a cross-section of cities. J Monet Econ 36:117–143. https://doi.org/10.1016/0304-3932(95)01206-2
Ishikawa Y (1992) The 1970s migration turnaround in Japan revisited: a shift share approach. Pap Reg Sci 71:153–173
Lucas RE (1988) On the mechanics of economic development. J Monet Econ 22:3–42. https://doi.org/10.1016/0304-3932(88)90168-7
Manson S, Schroeder J, Van Riper D, Ruggles S (2017) IPUMS national historical geographic information system: version 12.0 [Database]. University of Minnesota, Minneapolis. https://doi.org/10.18128/D050.V12.0
McHenry P (2014) The Geographic distribution of human capital: measurement of contributing mechanisms. J Reg Sci 54:215–248. https://doi.org/10.1111/jors.12067
Moretti E (2004) Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. J Econ 121:175–212. https://doi.org/10.1016/j.jeconom.2003.10.015
Moretti E (2012) The new geography of jobs. Houghton Mifflin Harcourt, New York
Plane D (1987) The geographic components of change in a migration system. Geogr Anal 19:283–289
Plane D, Heins F (2003) Age articulation of US inter-metropolitan migration flows. Ann Reg Sci 37:107–130
Plane DA, Rogerson PA (1991a) Tracking the baby boom, the baby bust, and the echo generations: how age composition regulates US migration. Prof Geogr 43:416–430. https://doi.org/10.1111/j.0033-0124.1991.00416.x
Plane DA, Rogerson PA (1991b) The ten commandments of migration research. In: Boyce DE, Nijkamp P, Shefer D (eds) Regional science: retrospect and prospect. Springer Berlin Heidelberg, Berlin/Heidelberg, pp 15–41. https://doi.org/10.1007/978-3-642-76311-3_2
Plane DA, Henrie CJ, Perry MJ (2005) Migration up and down the urban hierarchy and across the life course. Proc Natl Acad Sci U S A 102:15313–15318. https://doi.org/10.1073/pnas.0507312102
Romer PM (1986) Increasing returns and long-run growth. J Polit Econ 94:1002–1037. https://doi.org/10.1086/261420
Ruth M, Franklin RS (2014) Livability for all? Conceptual limits and practical implications. Appl Geogr 49:18–23. https://doi.org/10.1016/j.apgeog.2013.09.018
Storper M, Scott AJ (2009) Rethinking human capital, creativity and urban growth. J Econ Geogr 9:147–167. https://doi.org/10.1093/jeg/lbn052
U.S. Census Bureau (2006) Census 2000 PHC-T-41. A half-century of learning: historical statistics on educational attainment in the United States, 1940 to 2000
Walker KE (2017) The shifting destinations of Metropolitan migrants in the U.S., 2005–2011. Growth Chang 48:532–551. https://doi.org/10.1111/grow.12187
Whisler RL, Waldorf BS, Mulligan GF, Plane DA (2008) Quality of life and the migration of the college-educated: a life-course approach. Growth Chang 39:58–94. https://doi.org/10.1111/j.1468-2257.2007.00405.x

Chapter 6
Population and Employment Change in US Metropolitan Areas

Gordon F. Mulligan, Helena A. K. Nilsson, and John I. Carruthers

Abstract The regional adjustment model is used to analyze changes in population and employment across the American metropolitan landscape between 1990 and 2015. Estimates are made for the effects of natural and human-created amenities on population change and the effects of wages, self-employment, patents, economic specialization, and age composition on employment change. Short-run impacts, estimated by linear regression, allow identification of a 2 by 2 "growth operator" matrix; long-run impacts are estimated by powering this matrix. In the early years of the 25-year study period, employment numbers largely drove population change, but, once the direction of causality reversed, population numbers largely drove employment change. Clearly, the balance between the overall effects of population and employment can shift over long periods of time.

Keywords Regional adjustment model · Population change · Employment change · American metropolitan · Long-run impacts

G. F. Mulligan (✉)
School of Geography and Development, University of Arizona, Tucson, AZ, USA
e-mail: mulligan@email.arizona.edu

H. A. K. Nilsson
Institute of Retail Economics and Centre for Entrepreneurship and Spatial Economics, Jönköping International Business School, Jönköping, Sweden
e-mail: Helena.Nilsson@ju.se

J. I. Carruthers
Department of City and Regional Planning and Graduate Field of Regional Science, Cornell University, Ithaca, NY, USA
e-mail: john.carruthers@cornell.edu

R. S. Franklin (ed.), *Population, Place, and Spatial Interaction*, New Frontiers in Regional Science: Asian Perspectives 40,
https://doi.org/10.1007/978-981-13-9231-3_6

6.1 Introduction

Ever since the dramatic counterurbanization trends of the 1960s were noted by demographers and other social scientists, much interest has been focused on the so-called *chicken-or-egg* problem. Until that time, traditional thinking proposed that employment numbers were driving regional population change; in other words, analysts generally believed that people were following jobs. But, after the 1960s, these analysts increasingly noted that, in many parts of the nation, population numbers were instead driving regional employment change; in other words, jobs were often following people. But population and employment change, being interdependent, must co-exist everywhere even though it is difficult at times to determine which has the more important effect.

This chapter addresses the problem by using a regional adjustment model that allows current levels of population and employment to adapt to past levels of *both* population and employment, where the mutual adaptation is controlled by various place-specific conditions. These contextual variables influence how population numbers react to prior (or initial) natural and human-created amenities and how employment numbers react to prior wages, patent rates, self-employment, economic specialization, and the age composition of the workforce. Linear regression is used to estimate the four coefficients of a 2 by 2 "growth operator" matrix over a series of 10-year time periods, where these coefficients trace out the twin relationships connecting current population and employment numbers to their prior or lagged numbers. Matrix multiplication is then used to project, in 10-year increments, how these short-run relationships are expected to change in the future. Here, the column elements of the matrices reveal the balance existing between the overall population and employment effects, and, by repeated matrix multiplication, future shifts in this balance can be exposed. As matters turn out, this multiplication typically amplifies the relative importance of each effect as initially revealed in the growth operator matrix. The projected long-run relationships between population and employment numbers can prove to be very different from the estimated short-run relationships.

So, the main intent of the chapter is to shed light on the ever-shifting bidirectional relationships persisting between population and employment change among the metropolitan regions of the USA. However, the study has two other secondary intents that are worth noting at the outset. First, the short- and long-run estimates of the two overall effects are determined, in part, by the form (levels versus densities) of the input data, the time lag (10 years vs. 5 years) used in making the estimates, and whether the twin streams of the adjustment process are controlled for spatial dependency. Even given the limitations in space, the chapter sheds light on how each of these factors affects the estimated balance between the overall population and employment effects. Second, considerable research on entrepreneurship and innovation since the 1980s has outlined the main features of the so-called Knowledge Economy. The chapter sheds light on how activities like self-employment and patenting have affected the evolving relationships between population and employment across the metropolitan engines of the US space-economy. Evidence is given that the population and employment effects of the knowledge-rich and knowledge-poor areas have been somewhat different in the recent past.

The next section briefly summarizes the literature on the so-called chicken-or-egg problem in demography and migration studies. Here mention is made of hedonic models, which recognize that mobile households (and firms) can substitute high-amenity, low-wage locations for low-amenity, high-wage locations. The ensuing chapter then reviews the simplest regional adjustment models where population and employment respond to one another, over space and time, given an array of initial demographic and economic conditions. Next, ordinary least-squares (OLS) regression is used to estimate a series of adjustment models for 377 US metropolitan areas during 1990–2015. The base case considers changes in population and employment levels, uses a 10-year lag in the adjustment process, and does not consider spatial dependency. Regression estimates are made for the mutually adjusting population and employment numbers using four overlapping time periods: 1990–2000, 1995–2005, 2000–2010, and 2005–2015. Then consideration is given to how densities, as opposed to levels, will shift the various estimates in the base case and how uneven geographic nearness will affect those estimates as well. This part of the chapter ends with a more general perspective where the data are pooled across the four different time intervals. Here, alternative pooled estimates are also given for the series of nonoverlapping 5-year intervals between 1990 and 2015. Finally, the chapter closes with some suggestions regarding future directions for research.

6.2 Causality Between Population and Employment Change

Soon after viewing results from the 1970s Census of Population and Housing, Calvin Beale (1972) noted that many rural areas and peripheral regions of the nation were growing at the expense of the more urbanized and centrally located metropolitan areas. This novel turnaround process, although not permanent, gained the interest of demographers and other analysts who, given their disciplinary biases, saw it as an example of either employment restructuring or population deconcentration (Frey 1993). But, at nearly the same time, Richard Muth (1971) suggested that household migration in the USA should be studied as a chicken-or-egg problem because of the inherent uncertainty in the direction of causality between population and employment change. For quite some time, the conventional wisdom had been that people followed jobs, both within and between regions, which meant that employment was exogenous to the geographic distribution of population (Borts and Stein 1964). But a very different account of change in the space-economy slowly emerged where jobs often were seen to follow people, meaning that the twin distributions of population and employment had to be endogenously determined (Carruthers and Vias 2005). Radical as it might have seemed at the time, this bidirectional hypothesis is now widely accepted in the regional and social sciences.

The models devised to address bidirectional change involved ideas drawn from economists, demographers, planners, geographers, and others (Isserman 1986). One key stream of ideas arose from the work done on migration by people like Greenwood (1975) and Graves (1976) who noted that households often moved from places of high economic opportunity to places of low economic opportunity. This brought key demographic concepts, like the life cycle, to the forefront for consideration and testing (Graves 1979; Plane and Heins 2003). Here, Sjaastad (1962) proved especially influential because he suggested, somewhat earlier, that households look at long-distance migration as an investment decision. Another stream of ideas came from the path-breaking work done on hedonic markets by people like Rosen (1979) and Roback (1982), where it was argued that households might trade off higher wages and salaries for valued natural or human-created amenities. Yet another stream of research demonstrated that heterogeneity exists in mobility and migration choices, where households with very different attributes can make very different choices about where to live and work (Herzog and Schlottman 1986). Finally, other studies demonstrated that firms themselves could anticipate the preferences of their workers and locate, or even relocate, to areas that are rich in non-traded amenities (Boarnet 1994). There is now plenty of evidence, at least in the USA, that natural and human-created amenities elicit a steady effect on worker movements over fairly long periods of time while economic opportunity, typically more localized in space, affects worker movements in different ways at different points in time (Mueser and Graves 1995; Mulligan and Carruthers 2011). In any case, when assembled together, these various insights grant agency to both households and firms and mean, in support of Muth's original contention, that population and employment change should be simultaneously determined. From this perspective, the s*patial equilibrium* framework has become acceptable for explaining not only the short-term trends but also those long-term movements of households and firms that occur both within and between regions (Glaeser 2007).

It appears, then, that two very different growth processes are simultaneously unfolding in the more advanced space-economies. Following Bartik (1991) and DiPasquale and Wheaton (1996), these are usually labeled demand- and supply-induced growth. On the one hand, *demand-induced growth* occurs when firms expand employment, thereby causing an increase in the number of *jobs* in the regional labor market. On the other hand, *supply-induced growth* occurs when households relocate for choice, causing an increase in the number of *people* in the regional labor market. A major challenge to regional and social scientists is to identify those periods when economic opportunities, on the one hand, or personal preferences, on the other, have a greater impact on the ever-shifting population and employment numbers of the nation's many metropolitan areas.

6.3 The Adjustment Process

Regional adjustment models are spatial analogues to those adjustment models that have been widely adopted in the various branches of economics. Surprisingly, the spatial models originate from research on the distribution of people and jobs within regions (Steinnes and Fisher 1974; Steinnes 1977). But these spatial models became widely known only after the study by Carlino and Mills (1987), and then Clark and Murphy (1996) applied the adjustment framework to population and employment change across the continental counties of the USA during the 1970s and 1980s, respectively. Later studies addressed a variety of diagnostics and discussed issues like specification, scale, the effects of macroeconomic conditions, and the like (Mulligan et al. 1999; Hoogstra et al. 2017).

In all types of partial adjustment models, the variable of interest—usually population or employment—is seen to be constantly in motion but nevertheless moving toward some equilibrium position. Consequently, the current level of this variable is estimated by accounting for its past (lagged) level, the current level of the other variable it is adjusting to, and the lagged values of a number (vector) of other explanatory variables. So, in the simple 2 by 2 case, population adjusts to employment and, at the same time, employment adjusts to population. In practice, though, this means that estimates must first be made for both current population and current employment based on the lagged values for *both* variables. So, in the end, current population $POPUL_t$ is seen to adjust to an estimate for current employment $EMPLY*_t$, and, alternatively, current employment $EMPLY_t$ is seen to adjust to an estimate for current population $POPUL*_t$. Since the study by Carlino and Mills (1987), this adjustment period often makes use of Census data, so is usually assumed to be a decade in length, but the most appropriate time lag is not really known. Here, the pair of adjustment equations are estimated by two-stage least squares regression procedures where the second-stage results are:

$$POPUL_t = a_1 + b_1 POPUL_{t-1} + c_1 EMPLY*_t + \mathbf{d_1 VECTR_{t-1}} + e_1 \tag{6.1}$$

$$EMPLY_t = a_2 + b_2 POPUL*_t + c_2 EMPLY_{t-1} + \mathbf{d_2 VECTR_{t-1}} + e_2 \tag{6.2}$$

This means that the reduced forms of these two equations can be recovered by substituting for $EMPLY*_t$ in Eq. (6.1) and for $POPUL*_t$ in Eq. (6.2). When making the estimates for current employment and population, using both of their lagged values, it is customary to include all the other explanatory variables in $\mathbf{VECTR_{t-1}}$, else these estimates will likely be biased because the two distributions of errors are correlated. The reduced-form expressions are as follows:

$$POPUL_t = g_1 + h_1 POPUL_{t-1} + i_1 EMPLY_{t-1} + \mathbf{j_1 VECTR_{t-1}} + k_1 \tag{6.3}$$

$$EMPLY_t = g_2 + h_2 POPUL_{t-1} + i_2 EMPLY_{t-1} + \mathbf{j_2 VECTR_{t-1}} + k_2 \tag{6.4}$$

which can, of course, be estimated directly by OLS regression (or a similar technique). The coefficient c_1 (c_2) indicates the rate at which population (employment) is adjusting to employment (population), while both variables supposedly converge toward a spatial equilibrium. However, the possibility exists that the adjustment process might not reach this equilibrium if one or the other variable grows too quickly or too slowly. A test for convergence is therefore needed on the reduced-form equations (see below). Moreover, it is a simple matter to address changes instead of levels on the left-hand sides of Eqs. (6.3) and (6.4) and, in such cases, the estimates h_1 and h_2 become h_1-1 and h_2-1, respectively, while all the other estimates stay the same. More than two endogenous variables can be considered in the adjustment process and, in such cases, a third or fourth variable is sometimes chosen from those already included in the list $\mathbf{VECTR_{t-1}}$ of other explanatory variables. Several sources, including Mulligan and Nilsson (2020), show how actual numerical estimates are calculated on a step-by-step basis.

Tests exist to address stability or convergence. It is unclear, though, how important this theoretical property really is because so many other real-world factors—including shifts in fertility and mortality, changes in trade policy, and even outright international conflict—can disturb the twin paths of the adjustment process over time (Kaldor 1972). Nevertheless, it is customary to use the lagged coefficients of the 2 by 2 "growth operator" matrix $\mathbf{M} = (h_1, i_1; h_2, i_2)$, where the semicolon delimits the separate rows of that matrix (Keyfitz and Caswell 2005; Rogers 1971). Recall that this square matrix is comprised of the estimates found in the two reduced-form equations, where h represents population, i represents employment, and the subscripts signify the population and employment change equations, respectively. It is worth noting that while the sums of the estimates along the rows or down the columns of $\mathbf{M}$ often approximate unity, this is not a required property of the growth operator matrix in the present application. If real eigenvalues (characteristic roots) exist for this matrix, then convergence eventually takes place in the adjustment process, and in theory, an equilibrium exists. The dominant (larger) eigenvalue simply indicates the correct solution for stability in the adjustment process.

Convergence in the adjustment process allows specification of the so-called unit vector, which indicates the proportional or fractional importance of the two (or more) variables at the equilibrium. This property of the adjustment process is often overlooked in regional science although it certainly is informative in demography (Rogers 1968). In the current application, the unit vector is useful because it indicates whether the solution for the metropolitan labor markets is in fact sustainable over the long run. In the standard adjustment model, where regional employment is drawn solely from regional population, population should *at the very least* be equal to employment at the equilibrium. If the population fraction exceeds the employment fraction in the unit vector, then the adjustment process is sustainable; however, if the employment fraction exceeds the population fraction in the unit vector, then the process is unsustainable. This is the correct interpretation of the unit vector when analysts deal with large, stand-alone observation units like metropolitan areas. But when analysts deal with smaller spatial units that are contiguous, like census tracts, this rule must be relaxed in order to accommodate interaction across those units, including cross-commuting (Carruthers and Mulligan 2019). Also, when dealing with the mutual adjustment of variables other than population and employment numbers, the interpretation of the results can become more problematic.

However, stability can be examined in yet another way that is more useful in practice. When a theoretical equilibrium (denoted by an asterisk) does exist, the pattern of elements down each column of matrix $\mathbf{M}*$ is identical. In fact, the ratios between these various elements are the same as those down the unit vector. This means that the ratios between corresponding pairs of elements must be identical from one row of $\mathbf{M}*$ to the next, meaning that $h_1^*/i_1^* = h_2^*/i_2^*$ in the 2 by 2 case. The overall importance of each variable at the outset of the adjustment process can be determined by summing the various coefficients down each column of the growth operator matrix. This procedure is the same as that used to calculate the column multiplier in input-output analysis. By comparing those two sums, the analyst is given a short-run estimate of the separate population and employment effects, where each effect includes its own (on diagonal) and its cross (off diagonal) component. Repeated matrix multiplication (or squaring) can be used to determine how those two initial effects, in the absence of other forces, are expected to change on a round-by-round basis afterward. This procedure, in turn, is the same as that used by Markov models for population redistribution, although now the row coefficients do not necessarily sum to unity (Plane and Rogerson 1994). Here, as before, the analyst can sum the various coefficients down the columns of each "new" matrix and then compare those sums in order to update estimates of the overall population and employment effects on a round-by-round basis. To summarize, in the 2 by 2 case, the two overall effects are determined by the ratio $(h_1 + h_2) : (i_1 + i_2)$ in the short run, the ratio $(h_1 + h_2)_{r+1} : (i_1 + i_2)_{r+1}$ after round r ($r = 1, 2, \ldots$) in the long run, and the ratio $h_1^* : i_1^*$ or $h_2^* : i_2^*$ at the (theoretical) long-run equilibrium.

6.4 Variables and Conjectures

The analysis focuses on 377 of the 381 metropolitan statistical areas now monitored by the Bureau of Economic Analysis (BEA). Four cities in Alaska and Hawaii were omitted because they were extreme geographic outliers. In 1990, the mean population *POPUL* of these areas was approximately 246 K, but by 2015, this mean figure had risen to 319 K; in 1990, the mean total employment *EMPLY* of these areas was 132 K, but by 2015, this figure had risen to 182 K (BEA 2018). While many of the smaller places had not yet achieved metropolitan status by Census year 1990, by year 2000, many of these had at least achieved micropolitan status.

Besides prior population and current employment, current population was conjectured to be affected by the quality and quantity of various natural and human-created amenities. Here natural amenities were captured by both cooling degree-days *CDGDY*, which ranged from 109 to 3984, and heating degree-days *HDGDY*, which ranged from 245 to 9897. Both figures varied considerably by temperature, humidity, and moisture across the large land mass of the continental

USA (Savageau and Boyer 1993; Savageau 2007; BizEE degree days 2018). The climate measures were assumed to be constant over the 25-year study period, and adjustments were not made for any local variation in utility rates. Human amenities *HAMEN* were next estimated by first regressing median house values on per capita income, heating degree-days, and cooling degree-days, and then using the residuals as net measures of those house values. Based on current dollars, in 1990, the median house value averaged \$137 K, but by 2015, this figure had climbed to \$170 K; in 1990, average personal income was \$17.4 K, but by 2015, this figure had climbed to \$42.7 K (Savageau and Boyer 1993; U.S. Census Bureau 2018). The three conjectures were that population change would be driven lower by *CDGDY* (−) and *HDGDY* (−) but higher by *HAMEN* (+). The first two conjectures reflect the notion that households prefer mild to extreme climates and often seek out those locations offering either low cooling or heating degree-days. The second conjecture, based on the idea that human-created amenities are capitalized into higher house values, reflects the notion that households normally will pay for a vibrant local ambience and for those public goods that are highly valued like health and education services (Carruthers and Mundy 2006).

Besides prior employment and current population, current employment was conjectured to be significantly affected by average wages and salaries, industrial specialization, and patenting activity (see Mulligan and Nilsson 2020). Expressed in current dollars, annual wages averaged approximately \$20.7 K in 1990, but the average figure for *WAGES* rose to \$44.6 K by 2015 (BEA 2018). Although manufacturing jobs were considered in the earlier study by Mulligan and Nilsson (2020), industrial specialization *PPROF* was solely measured here by the high human-capital employment arising in the professional, scientific, and technical services (classified as NAICS 54). These knowledge-intensive jobs comprised 4.27% of all metropolitan jobs in 1990 and 5.18% in 2015 (BEA 2018). Patenting *PATEN* was included because this activity is known to differentiate between highly creative cities and less creative ones (Florida 2002; Mulligan et al. 2017). In 1990, the average patent density (per 1000 persons) was 0.161 across the 377 metropolitan economies, but later, in 2015, this figure had nearly doubled to 0.318 (U.S. Patent and Trade Office 2018). The first conjecture (*WAGES*, −) recognizes that firms generally prefer to pay lower wages to their workers, although this tendency varies a lot with industry and with worker productivity. The second conjecture (*PPROF*, +) indicates that, due to spillover and local learning effects, overall employment levels should increase more when technical and scientific jobs are initially high. The third conjecture *(PATEN*, +) recognizes that highly innovative metropolitan economies should generate more overall jobs than less innovative economies (Moretti 2012; Tsvetkova 2015; Mulligan 2018).

However, this study addresses two other conditions that were not included in the earlier study by Mulligan and Nilsson (2019). One of these is proprietary employment, which is a popular measure of the incidence of entrepreneurship in regional economies (Kirzner 1973; Godin et al. 2008). Self-employment can be measured in several different ways, but, for present purposes, the figures released in the BEA's Economic Profiles are used (BEA 2018). In 1990, self-employment *PROPR* comprised on average 15.7% of all metropolitan jobs, but by 2015, this proportion had steadily climbed to 20.5%. Finally, in order to control for the differential age composition of the various labor markets, a prime workforce variable *PWFOR* was calculated as the ratio between those persons in the 18–44 age cohorts and those persons in all age cohorts. People in the 18–44 age group are widely believed to be more productive, on average, than those in either the younger or older age groups. The mean of this prime workforce proportion fell from 31.4% in 1990 to 25.0% as the population aged in most metropolitan areas. Higher initial rates of self-employment (*PROPR*, +) and higher initial prime workforce ratios (*PWFOR*, +) were both conjectured to have a positive impact on overall job creation across the US metropolitan landscape during the 25-year study period.

6.5 Results

6.5.1 Regression Estimates

As mentioned earlier, the appropriate short-run estimates are the various reduced-form regression coefficients. Table 6.1 shows the first series of these estimates (the base case), using OLS regression procedures, where the study period has been divided into four overlapping decades. Here, the goodness-of-fit statistics, including the standard estimation error (SEE), indicate that the fit of the population equation is always superior to that of the employment equation; note, too, that this gap is greatest during the recessionary events of the 2000–2010 period. The five relevant estimates for population change are shown in the top panel, and the seven relevant estimates for employment change are shown in the bottom panel of the table. The other estimates (those without conjectures) in the two reduced-form equations are not shown. The SEEs also prove to be superior, in the range of 3.5–9.4%, to those found in the corresponding models developed earlier by Mulligan and Nilsson (2020), and this superiority generally widens over time during the 25-year study period. All the estimates are in logarithmic form, so the coefficients can be interpreted as elasticities.

Current population (or population change) was strongly driven by past population, but the direct effect of lagged employment was only significant in the first 10-year period. Human amenities and heating degree-days proved to be significant during the first three periods but not during the last. Here, heating degree-days generally had a stronger (negative) impact on population numbers than did cooling degree-days, indicating that the avoidance of cold weather—reflected in moves to places like Miami and Phoenix—was a much stronger determinant of population change than the avoidance of hot weather. However, current employment (or employment change) exhibited much more consistency. Here, the effect of lagged employment was always significant, as expected, but its coefficient declined during the full study period. However, the direct effect of lagged population proved to be significant only in the third period. Wages had a strong negative impact and both professional services, and self-employment had a strong positive impact in all four periods. The prime workforce ratio was significant in three of the four periods, but patenting activity was significant only in the first period. Although the pattern is not entirely clear, the results suggest that people following jobs was more important during the 1990s, while jobs following people became increasingly important afterward. This finding entirely concurs with the narrative about the chicken-or-egg problem found at the beginning of the chapter. As for the robustness of these results, the removal of *PWFOR* or *PROPR* from the regression equations did not shift any of the other estimates to a notable degree.

Various applications of the regional adjustment model have adopted population and employment densities instead of levels. Accordingly, Table 6.2 shows the reestimations for the four time periods using metropolitan densities that were generated from the data on land and water areas found in the 2017 Gazetteer Files (U.S. Census Bureau 2018). The goodness-of-fit statistics indicate that the two sets of estimates are comparable in their overall precision. The shifts made in the estimates for either lagged population or lagged employment proved to be very small, and in only one case, that of population change during 2000–2010, was there a change in the sign of a coefficient (one that is insignificant anyways). As for the role

Table 6.1 Reduced-form estimates: Levels and 10-year lags

	1990–2000	1995–2005	2000–2010	2005–2015
Population				
Constant	2.609*	1.167*	1.285*	0.489
POPUL	0.894*	0.945*	0.982*	0.965*
EMPLY	0.109*	0.050	0.018	0.029
HAMEN	0.083*	0.142*	0.135*	−0.010
CDGDY	−0.009	0.005	0.012	0.031*
HDGDY	−0.067*	−0.057*	−0.031*	−0.005
Ad. R-sq	0.993	0.994	0.995	0.994
SEE	0.089	0.081	0.072	0.087
Employment				
Constant	2.985*	1.914*	0.040	−0.384
POPUL	−0.020	0.039	0.141*	0.061
EMPLY	1.023*	0.955*	0.846*	0.939*
WAGES	−0.397*	−0.142*	−0.229*	−0.186*
PWFOR	0.290*	0.015	0.393*	0.341*
PROFS	0.040*	0.066*	0.135*	0.115*
PATEN	0.015**	−0.002	−0.006	0.010
PROPR	0.124*	0.117*	0.211*	0.200*
Ad. R-sq	0.993	0.994	0.990	0.991
SEE	0.091	0.086	0.112	0.108
Stable	Yes	Yes	Yes	Yes
Sustainable	No	Yes	Yes	No

Note: $n = 377$; * 0.01 level; ** 0.10 level

Table 6.2 Reduced-form estimates: Densities and 10-year lags

	1990–2000	1995–2005	2000–2010	2005–2015
Population				
Constant	1.779*	0.759	0.776**	0.239
POPUL	0.854*	0.932*	0.981*	0.961*
EMPLY	0.112*	0.044	−0.005	0.022
HAMEN	0.097*	0.148*	0.131*	−0.009
CDGDY	−0.012	0.002	0.012	0.030*
HDGDY	−0.079*	−0.064*	−0.038*	−0.008
Ad. R-sq	0.991	0.992	0.994	0.991
SEE	0.086	0.080	0.070	0.086
Employment				
Constant	2.246*	1.486*	−0.090	−0.781
POPUL	−0.055	0.024	0.139*	0.058
EMPLY	1.026*	0.949*	0.838*	0.927*
WAGES	−0.271*	−0.061	−0.189*	−0.130*
PWFOR	0.251*	−0.040	0.329*	0.332*
PROFS	0.045*	0.066*	0.128*	0.120*
PATEN	0.017**	0.001	−0.003	0.012
PROPR	0.089*	0.088*	0.188*	0.183*
Ad. R-sq	0.991	0.992	0.985	0.987
SEE	0.089	0.085	0.112	0.108
Stable	Yes	Yes	Yes	Yes
Sustainable	No	Yes	No	No

Note: $n = 377$; * 0.01 level; ** 0.10 level

of local conditions, there is again very little change that is evident in the coefficients of the population equation. However, the alternative specification does shift several coefficients of the employment equation: note that the elasticities (absolute values) for both wages and self-employment are consistently lower when densities are analyzed instead of levels. All in all, though, the two different specifications generate remarkably similar short-run estimates.

The earlier study by Mulligan and Nilsson (2020) indicated that spatial lags should probably be addressed in the estimation of the two adjustment equations. Consequently, the findings of Table 6.1 were revisited after accounting for the uneven spatial distribution of the nation's 377 metropolitan areas. Current population and employment levels were reestimated for each of the four 10-year intervals using a GS2SLS spatial lag model, where an inverse distance matrix was adopted, with a 400-kilometer threshold, so that every metropolitan area had at least one neighbor. Following Kelejian and Prucha (2010), a "minmax" normalized weight matrix (where each element is divided by the smallest of the largest column- and row-sum) was used in order to preserve the internal weighting structure. The new reduced-form results are shown in Table 6.3, where all eight of the estimates for spatial lags are once again negative. As before, this finding indicates that spatial

Table 6.3 Reduced-form estimates: Levels and 10-year lags with spatial dependence

	1990–00	1995–05	2000–10	2005–15
Population				
Constant	2.276*	1.024**	1.092*	0.320
POPUL	0.888*	0.950*	0.997*	0.974*
EMPLY	0.111*	0.044	0.002	0.019
HAMEN	0.093*	0.143*	0.133*	−0.009
CDGDY	−0.006	0.006	0.015**	0.033*
HDGDY	−0.060*	−0.054*	−0.026**	−0.001
SPLAG	−0.007*	−0.002	−0.004*	−0.002
Pseudo R-sq	0.993	0.994	0.996	0.994
Employment				
Constant	2.509*	1.600*	−0.193	−0.727
POPUL	−0.028	0.050	0.159*	0.080
EMPLY	1.027*	0.942*	0.827*	0.919*
WAGES	−0.296*	−0.089	−0.189*	−0.136**
PWFOR	0.173*	−0.038	0.350*	0.288*
PROFS	0.040*	0.062*	0.130*	0.110*
PATEN	0.017*	0.000	−0.004	0.012**
PROPR	0.081*	0.092*	0.187*	0.176*
SPLAG	−0.010*	−0.006*	−0.005**	−0.006**
Pseudo R-sq	0.994	0.994	0.990	0.991
Stable	Yes	Yes	Yes	Yes
Sustainable	No	Yes	Yes	No

Note: $n = 377$; * 0.01 level; ** 0.10 level

spillovers are not present at all, thereby suggesting that US metropolitan areas compete independently rather than cooperate interdependently for people and jobs. These spatial impacts are evidently the strongest in the employment equation where *SPLAG* proved to be significant in each of the four overlapping decades.

As was the case earlier, the effects of adding this new variable were more apparent in the employment equation than in the population equation. The introduction of a spatial lag consistently reduced the elasticities (absolute values) of *WAGES* and *PWFOR* in all four 10-year periods and reduced the elasticity of *PROPR* in the first three of those periods. Clearly, the property of geographic nearness had a greater impact on recent employment change than on recent population change across the American metropolitan landscape. But, it should be emphasized here that these conclusions are all based on global models, and that other approaches, like geographically weighted regression (GWR), are needed to discern how the property of spatial nearness differentially affects the results on a place-to-place basis.

In all three instances, the ever-changing adjustment process between population and employment leads to a stable solution across each of the four 10-year time periods. However, in many instances, that interaction is simply not sustainable, in a theoretical sense, because employment levels (or densities) eventually grow larger

Table 6.4 The two effects: 10-year lags and no spatial dependence

Round	Matrix coefficients	Column sums	Popul% Emply%
1990–2000			
1	0.894 0.109; −0.020 1.023	0.874 1.132	43.7 56.3
2	0.797 0.208; −0.038 1.044	0.759 1.252	37.8 62.2
3	0.708 0.300; −0.055 1.064	0.653 1.364	32.4 67.6
4	0.627 0.385; −0.071 1.083	0.556 1.466	27.5 72.5
16	0.366 0.657; −0.121 1.145	0.245 1.795	12.0 88.0
1995–2005			
1	0.945 0.050; 0.039 0.955	0.984 1.005	49.5 50.5
2	0.895 0.095; 0.074 0.914	0.969 1.009	48.9 51.1
3	0.849 0.135; 0.106 0.876	0.955 1.011	48.5 51.4
4	0.808 0.172; 0.134 0.842	0.942 1.014	48.1 51.9
16	0.675 0.284; 0.221 0.733	0.896 1.017	46.8 53.2
2000–2010			
1	0.982 0.018; 0.141 0.846	1.123 0.864	56.5 43.5
2	0.966 0.032; 0.258 0.718	1.224 0.750	62.0 48.0
3	0.953 0.044; 0.354 0.612	1.307 0.656	66.5 33.5
4	0.943 0.055; 0.434 0.524	1.377 0.579	70.4 29.6
16	0.913 0.081; 0.637 0.299	1.550 0.380	80.3 19.7
2005–2015			
1	0.965 0.029; 0.061 0.939	1.026 0.968	51.5 48.5
2	0.933 0.055; 0.116 0.883	1.049 0.938	52.8 47.2
3	0.904 0.078; 0.166 0.833	1.070 0.911	54.0 46.0
4	0.877 0.100; 0.211 0.786	1.088 0.886	55.1 44.9
16	0.790 0.166; 0.351 0.640	1.141 0.806	58.6 41.4

than population levels (or densities). But, as the next section reveals, it might in fact take many decades before those projected employment numbers come to exceed the population numbers.

6.5.2 Estimates of the Population and Employment Effects

As outlined earlier, the coefficients of the growth operator matrix provide the required estimates of the short-run (or immediate) population and employment effects. Recall that the elements are summed down the columns of matrix **M** so that, in the case of 1995–2005, the population total is 0.945 + 0.039 = 0.984 and the employment total is 0.050 + 0.955 = 1.005. Here, the grand total is 1.989, where the fractional contribution of population is 0.984/(0.984 + 1.005) = 49.5% and that of employment is 50.5% (see later).

The four period-specific estimates for these two effects are shown in Table 6.4 where each growth operator matrix is indicated on the top row as having a power of 1. This matrix is squared (raised to the second power) to provide estimates

10 years later where, in the example earlier, the elements of the "new" matrix in round two indicate the population and employment effects expected in 2015. After powering the matrix **M** 16 times (or squaring four times), the 4 coefficients typically take on a stable pattern, at least when the adjustment process converges. Note in Table 6.4 that the population and employment effects shift to 46.8% and 53.2%, respectively, once the estimates for 1995–2005 have gone through 16 rounds of matrix multiplication. These last two percentages differ slightly from those that would be expected at the (theoretical) long-run equilibrium, where calculations indicate that the two effects are 44.2% and 55.8%, respectively. As this example shows, at the equilibrium, it is entirely possible for the employment effect to exceed the population effect even when the projected population numbers in the (sustainable) unit vector exceed the projected employment numbers. However, these estimates are expressed in logarithms and, in order to arrive at the corresponding arithmetic figures, the two effects should be transformed by solving for exp.(0.442) and exp.(0.558), respectively, indicating that the arithmetic ratio is 47.1% and 52.9%. This transformation always reduces the ratios that have been estimated using logarithms.

Now, for each of the four decades, compare the various results that are shown along the bottom row in each case. The column sums for the growth operator matrix, calculated after 16 rounds of multiplication, indicate that the balance between the population and employment effects shifted a lot over the entire 25-year study period. In the first decade, the employment effect (88.0%) completely dominated the population effect (12.0%) but, by the third decade, the balance between those two effects had entirely reversed (80.3% vs. 19.7%). During the second and fourth decades, those two effects were approximately the same after 16 rounds of matrix multiplication. In all four instances, the initial gap between the two effects was amplified over time; in other words, if one effect was greater in the initial regression estimates, then the relative importance of that effect was monotonically increased in the matrix projections that ensued. Clearly, the estimates of the base case, which used a 10-year lag, suggest that "people followed jobs" early in the study period, but, sometime after Census year 2000, the trend shifted to one where "jobs followed people." Although this is speculation, the severe recessionary events experienced during the late 2000s might well have ended or dampened the second trend.

6.5.3 Some Comparative Results

A few more insights are gained by examining the findings shown in Table 6.5. Here, the estimates (bottom two cases) of this chapter are compared to those generated from the earlier model (top two estimates) developed by Mulligan and Nilsson (2020), where self-employment and the age of the workforce were not accounted for. Quite obviously, the population and employment effects of the updated model are more balanced in the sense that neither effect is quite so dominant (closer to 100%) in the long run. Spatial lags are included in both the second and fourth sets of estimates, and it seems that accounting for geographic nearness simply exaggerates

Table 6.5 Model specification and the two effects

	1990–2000		1995–2005		2000–2010		2005–2015	
SR Pop (MN)	0.961	0.042	1.023	−0.021	1.022	−0.014	0.938	0.058
SR Emp (MN)	−0.001	1.007	0.154	0.844	0.154	0.838	0.007	0.996
SR %	47.8	52.2	58.9	41.1	58.8	41.2	47.3	52.7
LR Pop	0.726	0.301	1.126	−0.108	1.079	−0.071	0.608	0.370
LR Emp	−0.007	1.056	0.790	0.207	0.779	0.210	0.044	0.978
LR %	34.6	65.4	95.1	4.9	93.0	7.0	32.6	67.4
SR Pop∗ (MN)	0.961	0.039	1.028	−0.029	1.034	−0.027	0.968	0.026
SR Emp∗ (MN)	−0.001	1.001	0.163	0.831	0.186	0.803	0.052	0.947
SR %	48.0	52.0	59.8	40.2	61.1	38.9	51.2	48.8
LR Pop	0.726	0.273	1.140	−0.145	1.195	−0.127	0.800	0.155
LR Emp	−0.007	1.007	0.813	0.157	0.877	0.105	0.310	0.676
LR %	36.0	64.0	99.4	0.6	100.1	−0.1	57.1	42.9
SR Pop	**0.894**	**0.109**	**0.945**	**0.050**	**0.982**	**0.018**	**0.965**	**0.029**
SR Emp	**−0.020**	**1.023**	**0.039**	**0.955**	**0.141**	**0.846**	**0.061**	**0.939**
SR %	**43.7**	**56.3**	**49.5**	**50.5**	**56.5**	**43.5**	**51.5**	**48.5**
LR Pop	**0.366**	**0.657**	**0.675**	**0.284**	**0.913**	**0.081**	**0.790**	**0.166**
LR Emp	**−0.121**	**1.145**	**0.221**	**0.733**	**0.637**	**0.299**	**0.351**	**0.640**
LR %	**12.0**	**88.0**	**46.8**	**53.2**	**80.3**	**19.7**	**58.6**	**41.4**
SR Pop∗	0.888	0.111	0.950	0.044	0.997	0.002	0.974	0.019
SR Emp∗	−0.028	1.027	0.050	0.942	0.159	0.827	0.080	0.919
SR %	43.0	57.0	50.4	49.6	58.2	41.8	52.9	47.1
LR Pop	0.328	0.664	0.708	0.243	0.982	0.009	0.842	0.105
LR Emp	−0.167	1.159	0.276	0.664	0.710	0.223	0.443	0.538
LR %	8.1	91.9	52.0	48.0	87.9	12.1	66.6	33.4

Note: MN refers to Mulligan and Nilsson (2020); SR denotes short run and LR denotes long run; SR and LR percentages are expressed in logarithms; ∗ includes spatial lag; boldface denotes base case

the gap in the two effects that exists in the nonspatial versions of the two models. In any case, the variety shown in the estimates reveals that the adjustment model provides estimates that are volatile, so extreme care must be taken when carrying out the appropriate regression estimations for the linked population and employment equations.

6.5.4 *Pooled Results*

Pooling the data was undertaken to provide more observations and to dampen some of the volatile results indicated earlier. The new estimates are shown in Table 6.6, where the 10-year time lags are compared to those arising from shorter 5-year time

Table 6.6 Pooled estimates: 10-year versus 5-year lags

	2nd stage	Reduced form	2nd stage	Reduced form
Population				
Constant	1.298*	1.362*	1.086*	1.163*
POPUL	0.941*	0.944*	0.889*	0.891*
EMPLY	0.059*	0.055*	0.111*	0.109*
HAMEN	0.065*	0.066*	0.042*	0.042*
CDGDY	0.008	0.009**	−0.002	−0.001
HDGDY	−0.040*	−0.041*	−0.029*	−0.030*
R-sq	0.994	0.994	0.998	0.998
SEE	0.084	0.085	0.051	0.051
Employment				
Constant	1.001*	1.083*	0.663*	0.688*
POPUL	0.060**	0.057**	0.021	0.019
EMPLY	0.940*	0.943*	0.980*	0.982*
WAGES	−0.227*	−0.238*	−0.124*	−0.126*
PWFOR	0.265*	0.277*	0.110*	0.111*
PROFS	0.074*	0.076*	0.038*	0.038*
PATEN	0.006	0.006	0.004**	0.004**
PROPR	0.163*	0.170*	0.091*	0.092*
R-sq	0.991	0.991	0.996	0.996
SEE	0.103	0.103	0.066	0.066
Stable	n.a.	Yes	n.a.	Yes
Sustainable	n.a.	No	n.a.	No

Note: n = 1508 (10 year), 1885 (5 year); * 0.01 level; ** 0.10 level; time dummies suppressed

lags. In both cases, the second-stage (Eqs. 6.1 and 6.2) and the reduced-form (Eqs. 6.3 and 6.4) estimates are shown. Spatial heterogeneity is not addressed here because the decade-specific results shown earlier were consistent, but three decade-specific time dummies were included to control for longitudinal effects.

In the base case, the highest elasticities are seen for *PWFOR* (0.277), *WAGES* (−0.238), and *PROPR* (0.170) in the reduced-form employment equation. So, holding other things constant, this suggests that a 1% increase in the rate of self-employment increased employment numbers by 0.170% in the average metropolitan economy. Also, human-created amenities appear to be slightly more important than natural amenities in the corresponding population equation, thereby supporting the contentions of Glaeser (2011), among others, about the importance of public goods and services in promoting healthy urban growth

However, halving the temporal lag down from 10 to 5 years has quite a dramatic impact on the various elasticity estimates. In fact, the importance of natural and human amenities is reduced by some 30%–35% in the population equation, and the importance of *PWFOR*, *WAGES*, *PROFS*, and *PROPR* is reduced by some 45%–60% in the employment equation. These results suggest that significant changes must take place in the 2 by 2 growth operator matrix, and indeed this proves to be the case.

With 10-year lags, the two column sums of **M** are nearly identical, 1.001 for population and 0.998 for employment, which suggests the two short-run effects are virtually the same. After 16 rounds of matrix multiplication, the population effect is 50.6%, and the employment effect is 49.4%; at the theoretical equilibrium, these two effects have converged on values of 50.7% and 49.3%. So, once the ups and downs in population and employment growth have been evened out with data pooling, the population numbers only have a marginally greater impact on overall metropolitan change than do the employment numbers. However, estimation with the shorter 5-year lags decreases the own effect for lagged population (from 0.944 to 0.891) in the first equation and increases the own effect for lagged employment (from 0.943 to 0.982) in the second equation. Moreover, the two cross effects are shifted a lot when adopting the shorter period of adjustment: the coefficient for *EMPLY* is doubled in the population equation, while that for *POPUL* is more than halved in the employment equation. This means the short-run estimate for the population effect is dampened, while that for employment is inflated, where the balance (expressed in logarithms) between the two effects becomes 45.5% versus 54.5% in the short run. After 16 rounds of matrix multiplication, the relative effects are 18.6% and 81.4%, while at the theoretical equilibrium, the twin effects converge on values of 14.7% and 85.3%, respectively. Once transformed into appropriate arithmetic terms, these proportions are population, 33.0%, and employment, 67.0%. So, the adoption of a shorter lag period in the estimation changes the balance between the two effects in favor of a dominant employment effect. Clearly, this is a crucial issue that deserves much more attention in studies that use the regional adjustment model.

6.6 Concluding Remarks

This chapter has shed new light on the issue of whether population or employment numbers have the most important effect on overall change in metropolitan areas. Here, regression-based short-run estimates, provided by the 2 by 2 regional adjustment model, have been used to generate corresponding long-run estimates. Overall change involves the total shifts in population and employment numbers experienced by regions, although some double-counting exists with the methodology. By adopting the nearly 400 metropolitan areas in the USA as observation units, the analysis reveals that "people followed jobs" more during the 1990s but "jobs followed people" more during the 2000s.

Future research might focus on improving various aspects of the estimation procedures. Other initial conditions, such as the taxes and expenditures of local governments, might improve the precision of both the population and employment equations. Moreover, data could be assembled that address the various birth and death rates of firms, or establishments, which is another indicator of local variation in entrepreneurship. Novel insights might also be gained by using other estimation procedures, including quartile regression or geographically weighted regression. The latter would seem to be especially promising as the effects of certain initial conditions might vary a lot across the metropolitan landscape. The findings might be of special interest to those analysts interested in the impacts of innovative or entrepreneurial behavior on metropolitan change.

In fact, the 2 by 2 approach can easily be expanded to a 3 by 3 approach by turning either patent volumes or self-employment numbers into endogenous variables. Preliminary estimates suggest that the results again prove to be stable although interpretation of the unit vector and the two separate effects becomes a bit more complicated. This more inclusive approach might also clarify, in part, whether it is more appropriate to use standard 10-year lags, or shorter 5-year lags, when estimating the population and employment equations in the regional adjustment model. A cursory investigation suggests that endogeneity might have different implications for change in the nation's smallest and largest metropolitan economies (Shearer et al. 2018).

References

Bartik T (1991) Who benefits from state and local economic development policies? W.E. Upjohn Institute, Kalamazoo

Beale C (1972) Rural and nonmetropolitan population trends of significance to National Population Policy. Economic Research Service, Washington, DC

BizEE Degree Days (2018) See https://www.degreedays.net. Last accessed 11 Jan 18

Boarnet M (1994) An empirical model of intrametropolitan population and employment growth. Pap Reg Sci 73:135–152

Borts G, Stein J (1964) Economic growth in a free market. Columbia University, New York

Bureau of Economic Analysis (2018) Interactive tables: personal income and employment. See https://www.bea.gov/regional/index/htm. Last accessed 11 Jan 18

Carlino G, Mills E (1987) The determinants of county growth. J Reg Sci 27:135–152

Carruthers J, Mulligan G (2019) The regional adjustment model: an instrument of evidence-based policy. Forthcoming in: Capello R, Nijkamp P (eds) Handbook of regional growth and development theories, revised and 2nd edn. Edward Elgar, Cheltenham, UK

Carruthers J, Mundy B (eds) (2006) Environmental Valuation. Ashgate, Burlington

Carruthers J, Vias A (2005) Urban, suburban, and exurban sprawl in the Rocky Mountain West: evidence from regional adjustment models. J Reg Sci 45:21–48

Clark D, Murphy C (1996) Countywide employment and population growth. J Reg Sci 36:235–256

DiPasquale D, Wheaton W (1996) Urban economics and real estate markets. Prentice-Hall, Englewood Cliffs

Florida R (2002) The rise of the creative class. Basic Books, New York

Frey W (1993) The new urban revival in the United States. Urban Stud 30:741–774

Glaeser E (2007) The economics approach to cities. Working paper 13696. National Bureau of Economic Research, Cambridge, MA

Glaeser E (2011) Triumph of the City. Penguin Press, New York

Godin K, Clemens J, Veldhus N (2008) Measuring entrepreneurship: conceptual frameworks and empirical indicators. Studies in entrepreneurship and markets. Fraser Institute, Vancouver

Graves P (1976) A re-examination of migration, economic opportunity, and the quality of life. J Reg Sci 16:107–112

Graves P (1979) A life-cycle empirical analysis of migration and climate, by race. J Urban Econ 6:135–147

Greenwood M (1975) Research on internal migration in the United States: a survey. J Econ Lit 13:397–433
Herzog H, Schlottman A (1986) What can be learned from the recent migrants? Growth Chang 17:37–50
Hoogstra G, van Dijk J, Florax R (2017) Do jobs follow people or people follow jobs? A meta-analysis of Carlino-Mills studies. Spat Econ Anal 4:357–378
Isserman A (ed) (1986) Population change and the economy: social science theories and models. Kluwer-Nijhoff, Boston
Kaldor N (1972) The irrelevance of equilibrium economics. Econ J 82:1237–1255
Kelejian H, Prucha I (2010) Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. J Econ 157:53–67
Keyfitz N, Caswell H (2005) Applied mathematical demography, 3rd edn. Springer, New York
Kirzner I (1973) Competition and entrepreneurship. University of Chicago, Chicago
Moretti E (2012) The new geography of jobs. Houghton Mifflin Harcourt, New York
Mueser P, Graves P (1995) Examining the role of economic opportunity and amenities in explaining population redistribution. J Urban Econ 37:176–200
Mulligan G (2018) Patent generation in U.S. metropolitan areas. In: Bernhard I (ed) Revised papers presented at the 21st Uddevalla symposium on innovation, entrepreneurship and industrial dynamics in internationalized regional economies, Luleå, Sweden, 14–16 June
Mulligan G, Carruthers J (2011) Amenities, quality of life, and regional development. In: Marans R, Stimson R (eds) Investigating quality of urban life. Springer, Dordrecht, pp 107–133
Mulligan G, Nilsson H (2020) Recent population and employment change in U.S. metropolitan areas. Forthcoming in: Chen Z, Bowen W, Whittington D (eds) Development studies in regional science: essays in honor of Kingsley E. Haynes. Springer, Tokyo
Mulligan G, Vias A, Glavac S (1999) Initial diagnostics of a regional adjustment model. Environ Plan A 31:855–876
Mulligan G, Reid N, Carruthers J, Lehnert M (2017) Exploring innovation gaps in the American space economy. In: Jackson R, Schaeffer P (eds) Regional research frontiers—vol. 1. Springer, Dordrecht, pp 21–50
Muth R (1971) Migration: chicken or egg? South Econ J 37:295–306
Plane D, Heins F (2003) Age articulation of U.S. inter-metropolitan migration flows. Ann Reg Sci 37:107–130
Plane D, Rogerson P (1994) The geographical analysis of population. John Wiley and Sons, New York
Roback J (1982) Wages, rent, and the quality of life. J Polit Econ 90:1257–1278
Rogers A (1968) Matrix analysis of interregional population growth and distribution. University of California Press, Berkeley
Rogers A (1971) Matrix methods in urban and regional analysis. Holden-Day, San Francisco
Rosen S (1979) Wage-based indexes of urban quality of life. In: Mieszkowski P, Straszheim M (eds) Current issues in urban economics. Johns Hopkins University Press, Baltimore, pp 74–104
Savageau D (2007) Places rated almanac, 7th edn. Places Rated Books, Washington, DC
Savageau D, Boyer R (1993) Places rated almanac, 4th edn. Macmillan, New York
Shearer C, Shah I, Friedhoff A, Berube A (2018) Metro monitor 2018. Brookings Institute. See https://www.brookings.edu/research/metro-monitor-2018. Last accessed 11 Jan 18
Sjaastad L (1962) The costs and returns of human migration. J Polit Econ 70(Supplement):80–93
Steinnes D (1977) Causality and intraurban location. J Urban Econ 4:69–79
Steinnes D, Fisher W (1974) An econometric model of intraurban location. J Reg Sci 14:65–80
Tsvetkova A (2015) Innovation, entrepreneurship, and metropolitan economic performance: empirical test of recent theoretical propositions. Econ Dev Q 29:299–316
U.S. Census Bureau (2018) See https://www.census.gov. Last accessed 11 Jan 18
U.S. Patent and Trade Office (2018) Calendar year patent statistics: listing of all U.S. metropolitan and micropolitan areas, total patent utility counts, 2000–2015. See https://www.uspto.gov/web/offices/ac/ido/oeip/taf/reports_cbsa.htm. Last accessed 11 Jan 18

Chapter 7
Confronting Statistical Uncertainty in Rural America: Toward More Certain Data-Driven Policymaking Using American Community Survey (ACS) Data

Jason R. Jurjevich

Abstract Aging and lacking infrastructure are major impediments to economic development in rural America. To address these issues, civic leaders often look to state and federal infrastructure grant/loan funding, where eligibility is often based on income requirements established by the US Census Bureau's American Community Survey (ACS). The problem, especially for rural communities, is that ACS data contain a high degree of statistical uncertainty (i.e., margin of error) that is often disregarded for determining program eligibility. For rural communities with unreliable income estimates, the most common work-around involves hiring a consultant to conduct an income census or survey to formally challenge the US Census Bureau's ACS estimate. Many rural communities, however, elect not to formally challenge unreliable ACS estimates either because they are unaware that reimbursement for conducting an income survey is an allowable expense under some grant/loan programs or they are dissuaded by the necessary time and resources. First, I summarize whether federal infrastructure grant/loan programs incorporate MOE values when determining community eligibility. Second, I examine the degree to which ACS estimates are statistically reliable for communities across rural America. Finally, using an example from Oregon, I recommend guidelines for how states can assist rural communities with statistically unreliable ACS estimates. These findings can help rural communities secure infrastructure funding that advances economic development and quality of life, and potentially support reliable data-driven policy and decision-making more broadly.

Keywords American Community Survey (ACS) · Rural · Demographic data · Margin of error (MOE) · Data-driven policy

J. R. Jurjevich (✉)
Toulan School of Urban Studies and Planning, Population Research Center (PRC), Portland State University, Portland, OR, USA
e-mail: jjason@pdx.edu

R. S. Franklin (ed.), *Population, Place, and Spatial Interaction*, New Frontiers in Regional Science: Asian Perspectives 40,
https://doi.org/10.1007/978-981-13-9231-3_7

7.1 Introduction

Rural communities across the USA are facing similar developmental challenges. The lack of water, sewerage, and transportation infrastructure, combined with rapidly aging assets more broadly, are major impediments to advancing economic development. Rural civic leaders often look to secure funding from state and federal infrastructure grant/loan programs to address these issues. Qualifying for grants/loans typically requires that the community meet strict requirements around socioeconomic measures that are often determined using the US Census Bureau's (USCB) American Community Survey (ACS). In 2008, ACS data determined 29% ($416 billion) of government assistance program funding and 69% ($389 billion) of all federal grant funding (Carpenter and Reamer 2010).

Unlike the decennial census, an enumeration of an entire population, ACS data are drawn from a statistical sample and therefore contain sampling error. As with all survey data, sampling error represents the risk that data drawn from statistical samples may not accurately represent the broader population. The Bureau conveys this risk, referred to as statistical uncertainty, through a margin of error (MOE) statistic that accompanies ACS estimates.

Navigating the challenges of statistical uncertainty in ACS data and other sources of quantitative demographic data is a key challenge in today's data-driven world. Overcoming these challenges is even more difficult for rural communities, where MOE values are often disregarded for determining federal program eligibility (Nesse and Rahe 2015). Consider a town of 2000 people in rural Lake County, Oregon. The town's median household income (MHI) estimate, according to 5-year ACS data, is $35,900 with an accompanying MOE value of ±$3500. Given that most grant/loan programs disregard MOE values, this town would be considered ineligible because the MHI figure exceeds the $35,000 grant threshold. This approach is problematic because when the MOE is considered, the estimate range is $32,400 ($35,900 − $3500) to $39,400 ($35,900 + $3500): there is a roughly equal chance that the town is in fact eligible.

What does a rural community do in a situation like this? The typical response is to pursue alternative approaches that establish the "true" population value in order to qualify for grants/loans. The most common work-around involves hiring a consultant to conduct an income census or survey to formally challenge the USCB's ACS estimate. Many rural communities, however, elect not to pursue this approach. Often, either they are unaware that some grant/loan programs reimburse costs for conducting an income census or they are dissuaded by the necessary time and resources.

Challenges that arise from statistical uncertainty introduce state-level policy implications and yield real, on-the-ground effects. Currently, states must prioritize financial resources to ensure that eligible communities receive their share of federal funding for community development. This challenge is particularly difficult for states with limited financial resources and/or a large rural population. Without sufficient funds to file a formal challenge, towns often delay or abandon capital projects until they can secure resources to hire a consultant. This situation creates negative impacts on citizens, particularly among the most vulnerable—including racial/ethnic minorities and low-income populations—by forcing individuals to commute long distances for affordable housing, food, and work.

Best practices for working with ACS data and practical examples and alternative approaches for reducing MOE values are well-documented in the existing literature. Aggregating geographies to reduce MOE, for example, is generally not a feasible approach because many infrastructure grant/loan programs require rural communities use place-level income estimates from the ACS. To highlight the challenge of statistical uncertainty in rural America and flesh out the unique challenge for rural communities—there are few, if any, alternative approaches to improving data quality—I address two interrelated questions in this chapter: (1) To what extent are ACS income estimates unreliable for rural communities across the USA? and (2) What policies can states implement to assist rural communities with statistically unreliable ACS estimates and qualify them for infrastructure grant/loan programs?

This chapter proceeds as follows: first, I summarize whether federal infrastructure grant/loan programs incorporate MOE values when determining community eligibility. Second, I examine the degree to which ACS estimates are statistically reliable for communities across rural America. Finally, using an example from Oregon, I recommend approaches and policies that states can implement to help rural communities with statistically unreliable ACS estimates. Presently, MOE values are often disregarded for determining program eligibility, and state support is either insufficient or nonexistent to help rural communities formally challenge unreliable ACS estimates. Given this scenario, I offer recommendations and brightlines for prioritizing state resources to account for MOE values when determining infrastructure grant/loan eligibility. These results are critical steps toward more certain data-driven policy and decision-making in rural America.

7.2 American Community Survey (ACS)

Between 1970 and 2000, the USCB administered two different forms to collect decennial census data: the short and long form (USCB 2019). In Census 2000, five out of six households received the short form, which contained approximately ten questions gathering data on age, sex, race/ethnicity, household relationship, and housing tenure. The long form surveyed one of six households and asked more detailed social, economic, and housing-specific questions. Long-form data were obsolete soon after each decennial census, however, a significant limitation for rapidly changing communities (Citro and Kalton 2007).

Rising costs of administering the long form, demand for timelier sociodemographic data, and concerns around confidentiality led the USCB to implement the ACS in 2005, replacing the long form (Torrieri 2007). ACS data, like

$$CV = \frac{SE}{\hat{X}} \times 100$$

Fig. 7.1 Coefficient of variation CV

decennial long-form data, are drawn from a statistical sample and are an approximated quantity rather than actual, true counts. This means that the degree of statistical uncertainty, expressed in the MOE statistic, represents a range of values expected to contain the true value of the quantity being estimated. For example, when the MHI estimate and corresponding MOE ($35,900, ±$3500) for a town in Lake County, Oregon, are considered, the range of values containing the true estimate is $32,400 ($35,900 − $3500) to $39,400 ($35,900 + $3500). The USCB reports MOE values with 90% statistical confidence, meaning that users desiring greater statistical confidence (e.g., 99%) will encounter greater statistical uncertainty.

Compared to long-form data, ACS data are drawn from a greatly reduced sample size. To put this into context, consider that the long form was administered to 1 in 6 households, while ACS estimates are derived from a sample of roughly 1 in 40 households.[1] The differences in sample size are meaningful for three key reasons. First, because long-form data were drawn from such a large statistical sample, the USCB did not report MOE values for long-form estimates (Spielman et al. 2014). This has, in part, created a situation where some data users do not understand statistical uncertainty in ACS data and avoid engaging with or reporting MOE values altogether (Jurjevich et al. 2018). Second, ACS data contain a greater degree of statistical uncertainty than decennial long-form data. USCB officials initially estimated that ACS estimates would have a 33% higher sampling error than long-form estimates. In the end, however, the sampling error ended up being roughly 75% higher than decennial long-form estimates (Spielman et al. 2014; Navarro 2012). Third, MOE values are higher for cross-tabulated data (e.g., child poverty) and for small-area and rural geographies.

To reduce statistical uncertainty and corresponding MOE values in ACS data, scholars recommend various alternative approaches, including collapsing data detail or aggregating census geographies (Citro and Kalton 2007; Heuvelink and Burrough 2002; National Academy of Sciences 2015; Spielman and Folch 2015; Spielman et al. 2014; USCB 2009). However, these recommendations are not feasible strategies for rural communities trying to become eligible for state and federal infrastructure grant/loan programs because many programs require rural communities use place-level income estimates from the ACS.

[1]Sampling odds are an effective way to compare the sampling frame of the ACS to the decennial long form. However, to be clear, the ACS sample is drawn from housing units and the group quarters population and is not based on a sampling rate. The 2017 ACS sample, for example, was drawn from roughly 2.1 million housing units and almost 300,000 individuals rising in group quarters, which together include more than 5,000,000 individuals.

A common way to express survey data reliability is through a statistic called the coefficient of variation (CV) (Fig. 7.1). Recommended by the USCB to evaluate data reliability, the CV expresses sampling error relative to the estimate and is calculated by dividing the standard error (SE) by the statistical estimate and multiplying the result by 100 (USCB 2009). Larger CV values indicate lower reliability. The Environmental Systems Research Institute's (ESRI 2014) "red-yellow-green" schematic is most often cited because the color schematic, combined with the simplicity of the CV values, is useful for conveying the reliability of ACS data to users unfamiliar with the concepts of data reliability (Jurjevich et al. 2018). Here, CV values less than 12% indicate a high degree (i.e., green) of reliability; CVs between 12 and 40 are somewhat (i.e., yellow) reliable; and CVs greater than 40 indicate little, if any (i.e., red), reliability.

Consider, for example, the estimate of children living at or below the poverty level is 20% (±3%) for Tract A and 25% (±15%) for Tract B. After deriving the standard errors from the MOE values (1.8% for Tract A and 9.1% for Tract B), the CVs are 9.0% and 36.4% for the two tracts, respectively.[2] This means that the poverty estimate for Tract 1 is reliable while the estimate for Tract 2 is somewhat reliable.

One key limitation of CV values is they are somewhat subjective; there are no hard-and-fast rules on cutoff values. To illustrate this point, consider these two different standards: (1) according to the National Academy of Sciences (2015), the USCB judges an estimate to be statistically reliable when the tract-level estimate for a "key variable" is less than or equal to 30%, and (2) the National Research Council (see Citro and Kalton 2007), on the other hand, recommends CV values be less than or equal to 10% as a standard level of precision. Together, these differences underscore a critically important point: achieving consensus on generally accepted CV threshold values for statistical reliability (or for specific use-case scenarios) is essential for professionals working in government, academia, and applied practice, and can also make it easier for novice data users to navigate and account for statistical uncertainty in ACS data.

7.2.1 Sampling Methodology

To ensure that ACS data are representative across age, sex, race/ethnicity, educational attainment, and urban/rural communities, the USCB implements a highly complex sampling methodology. What is particularly salient is the way the Bureau develops and implements its sampling methodology for targeting households. The likelihood of being selected for the ACS sample varies according to the geography

[2] The standard error is calculated by dividing the MOE values (3 and 15%) by 1.645 (z-score at 90% statistical confidence, which is the statistical level of confidence at which the USCB reports the MOE value).

Table 7.1 Corresponding CV values of ACS sampling methodology, 2005–2010

Tract size category	Average tract size	Coefficient of variation (CV)
0–400	291	66%
401–1000	766	41%
1001–2000	1485	29%
2001–4000	2636	26%
4001–6000	4684	19%
6000 +	8337	15%

Source: Asiala (2012)

Table 7.2 Corresponding CV values of ACS sampling methodology, 2011–present

Tract size category	Average tract size	Coefficient of variation (CV)
0–400	291	41%
401–1000	766	30%
1001–2000	1485	29%
2001–4000	2636	29%
4001–6000	4684	29%
6000 +	8337	28%

Source: Asiala (2012)

(i.e., urban or rural area) and the size of the census tract (i.e., the deviation from the average tract size of 6000 residents). For each ACS executed between 2005 and 2010, the sample was largely drawn from larger census tracts in order to yield more reliable estimates (Asiala 2012). As such, Table 7.1 shows that CV values were lowest and most reliable for larger (more populated) census tracts, which more often than not are in urban areas. On the other hand, smaller census tracts (largely in rural areas) contained higher CV values, indicating lower levels of statistical reliability.

In 2011, the Bureau made two significant changes in the ACS sampling methodology: (1) increasing the sample size from 2.9 to 3.5 million housing units and (2) changing the sampling frame to select more cases from smaller tracts, largely in rural areas (see National Academy of Sciences 2015, Alvarez and Salvo 2014, and USCB 2011). Table 7.2 illustrates how the combination of a larger ACS sample size, along with the shift in methodology, resulted in more equitable reliability across different sized census tracts. The main takeaway for rural areas is that data from the 2011 ACS (and subsequent years) should yield more reliable estimates than earlier ACS periods (i.e., 2010 and prior).

7.3 Federal Grant and Loan Infrastructure Programs

Many federal grant and loan programs rely on ACS data to determine program eligibility, to allocate funds, and for program assessment. Federal agency programs using ACS income data include, but are not limited to, the following: the US Department of Housing and Urban Development's (HUD) assisted housing programs; identifying distressed or underserved nonmetropolitan areas as part of the Community Reinvestment Act (CRA); determining the grant amount and fixed interest rate for community loans related to the US Department of Agriculture (USDA) Rural Development Wastewater Infrastructure Fund; setting eligibility parameters for the US Department of Agriculture Supplemental Nutrition Assistance Program (SNAP); and determining which communities are eligible for Technical Assistance and Public Works grants from the US Economic Development Administration.

Despite their widespread use, many agencies use ACS income data to determine program eligibility without accounting for accompanying MOE values. Nesse and Rahe (2015) explored this issue in more depth. They specifically looked at how the US Department of Housing and Urban Development (HUD), Department of Agriculture, and the Department of Transportation[3] use ACS data for administering their programs. Of particular interest was how these agencies incorporated MOE, if at all, in policy governance. Nesse and Rahe (2015) found that although most federal agencies recognize that ACS estimates are subject to accompanying statistical error, they generally do not consider MOE because the added complexity provides a marginal program benefit. Along these lines, the USDA Rural Development issued an administrative notice in 2012 instructing state and local officials to administer loan guarantee and grant programs using income and poverty data from the ACS. The notice, however, does not mention or establish policies for incorporating MOE (USDA 2012).

One of the only federal agencies to explicitly consider statistical uncertainty of ACS data in their programs, particularly for small areas, is HUD.[4] According to Usowski et al. (2008, p. 206), HUD aims to "minimize the possibility of publishing income estimates in which the annual change is more a reflection of the variation in estimation errors than a reflection of changes in underlying economic conditions."

[3]The specific programs examined by Nesse and Rahe (2015) include the Housing Choice Voucher Program, Supplemental Nutrition Assistance Program (SNAP), and the Urbanized Area Formula Program.

[4]Although HUD considers MOE for developing annual income limits, the agency does not consistently report MOE across all data programs. For example, see the Comprehensive Housing Affordability Strategy (CHAS) dataset at: http://www.huduser.org/portal/datasets/cp.html

To this end, HUD develops annual income limit estimates[5] based on a calculation that gives less emphasis to area estimates with high MOE values.

Why, given the potentially serious limitations, do some grant/loan programs disregard statistical uncertainty in their eligibility criteria? In addition to Nesse and Rahe's (2015) conclusion that agencies generally do not consider MOE, another possible factor is there may be little incentive to address this problem. Underscoring this point, in 2017 the Appropriations Committee in the US House of Representatives approved a measure mandating that HUD report areas in the USA where income data from the ACS, used to determine program eligibility, had an accompanying MOE of 20% or higher (Caster 2017). The provision, part of the FY 2018 omnibus spending package, was approved in March 2018.[6]

7.4 Geographic Delineation of "Rural America"

The geographic delineation of areas considered "rural" has long been of interest to researchers. As a result, there is a well-established body of research that defines, quantifies, and operationalizes varying degrees of rurality (e.g., Isserman 2005, Cromartie and Swanson 1996, McGranahan et al. 1986). A common starting point is the USCB's urban-rural classification. With census blocks as the primary geographic unit, the Bureau uses population density, land use, and distance measures to determine whether a census block qualifies as urban (Ratcliffe et al. 2016). Urban areas include communities that meet one of two following subclassifications: (1) an urbanized area (UA), which includes 50,000 or more people, and (2) an urban cluster (UC), which includes at least 2500 and fewer than 50,000 people.[7] Population, housing, and areas outside of urbanized areas and urban clusters are considered rural (Ratcliffe et al. 2016).

A limitation of the USCB's urban-rural continuum is that the binary nature of the definition is not particularly well-suited for classifying small towns. Small towns of 3000 to 5000, for example, are often classified as urban.[8] Consider the village of Lancaster, Wisconsin, a community of almost 4000 located in a remote corner of

[5]As Usowski et al. (2008) point out, annual income limits determine eligibility for the Public Housing program, Section 8 Housing Assistance Payments program, Section 2020 Supportive Housing for the Elderly, and Section 811 Supportive Housing for Persons with Disabilities.

[6]In response to the directive, HUD now publishes the MOE data for all block groups and all places. See: https://www.hudexchange.info/programs/acs-low-mod-summary-data/

[7]Ratcliffe et al. (2016) note that a minimum of 1500 people must reside outside of group quarters in order for an area to be classified as urban.

[8]Scholars have developed innovative approaches, including the ERC Rural-Urban continuum codes, that contextualize the degree of rurality.

Southwest Wisconsin. In 2010, the USCB classified Lancaster as urban.[9] Arguably, communities like Lancaster are more closely aligned with small towns of a few hundred people compared to more urbanized places. Given this situation, I define rural America by including rural places and small towns (possibly considered urban by the USCB) with fewer than 20,000 people.[10]

More populated small towns (e.g., a community of 16,500) might challenge traditional notions of what constitutes "rural." However, these communities are included for a practical reason. All communities with populations less than 20,000—rural places *and* small towns alike—do not have access to single-year annual ACS data (e.g., 2017) for data-driven decision-making. Instead, they must rely on a 5-year combined ACS data (e.g., 2013–2017). In the end, the 20,000 population threshold is an appropriate cutoff for revealing the challenges that both rural and small-town communities face: navigating the statistical uncertainty of ACS data for data-driven policymaking.

To assess the statistical reliability of ACS estimates across rural America, I selected a state from each of the USCB-defined regions (i.e., Northeast, South, Midwest, and West) (see Appendix). An upside to this approach is that it also makes it possible to assess the degree to which, if any, statistical reliability of ACS data varies across rural America by census region.

More than one in four Americans (28.8%) lived in a rural area or urban cluster in 2010 (Table 7.3). In the Midwest and South, more than one-third of residents, 37.2% and 33.8%, respectively, lived in a rural area or urban cluster in 2010. Across the West and Northeast, just 19.4% and 20.3% of residents, respectively, lived in rural areas or urban clusters. After examining data for all 50 states, I selected New Hampshire, North Carolina, Oregon, and Wisconsin as representative states for their respective census regions (Table 7.3). The population residing in rural and urban clusters in these states was slightly above each region's median.

In this analysis I rely on median household income (MHI) to assess the degree to which ACS estimates are statistically reliable. The decision is based on two key factors. First, MHI is used by most state and federal infrastructure grant/loan managers more for determining program eligibility. Second, MHI applies to all households and is not specific to a particular subpopulation. This is important because cross-tabulated data (e.g., child poverty) are drawn from a smaller population (i.e., children), resulting in higher MOE values. Therefore, although the challenges of statistical uncertainty of ACS data are more severe for cross-tabulated data, using MHI most closely illustrates the statistical challenges that rural and small-town communities must confront to secure infrastructure grant/loan funding.

[9]According to Census 2010 data, the population of Lancaster, WI was 3868. Roughly 94% (3642 persons) of the population was classified as urban and the remaining 6% (226 persons) was classified as rural (US Census 2010a).

[10]The mean and median CV values calculated for places with fewer than 20,000 residents include both incorporated towns and cities (e.g., Drain, OR), as well as for census-designated places (CDPs) (e.g., Glide, OR). This analysis excludes instances where MHI estimates are unavailable for a place/CDP.

Table 7.3 Urban-rural population classification for selected states and USA, 2010

	Total population	Urban population	Urban population, urbanized areas (UA)	Urban population, urban clusters (UC)	Rural population	Population share, rural and urban cluster (UC)
New Hampshire	1,316,470	793,872	623,168	170,704	522,598	52.7%
North Carolina	9,535,483	6,301,756	5,232,799	1,068,957	3,233,727	45.1%
Oregon	3,831,074	3,104,382	2,393,393	710,989	726,692	37.5%
Wisconsin	5,686,986	3,989,638	3,173,382	816,256	1,697,348	44.2%
USA	308,745,538	249,253,271	219,922,123	29,331,148	59,492,267	28.8%

Source: US Census Bureau (2010a)

7.5 Statistical Uncertainty in Rural America

The level of statistical uncertainty for MHI estimates among communities with less than 20,000 people is shown in Table 7.4. Expressed through mean and median descriptive statistics, the data illustrate three important points. First, MHI estimates became more reliable for rural and small-town communities during the two periods, 2006–2010 and 2013–2017 (i.e., median CV declined from 13.6 to 12.4). This is principally due to the change in ACS sampling strategy, which effectively "borrowed strength" from more populated urbanized areas to improve statistical reliability for less populated areas. Second, the increased reliability in MHI estimates for rural and small-town communities varied by state. Across rural and small towns in New Hampshire, for example, the reliability of MHI estimates improved the most (i.e., median CV from 18.0 to 13.7), while reliability remained largely constant for Wisconsin communities (i.e., median CV from 9.7 to 9.6). Third, although MHI estimates have improved over the past decade, the median level of statistical uncertainty for the most current MHI estimates remains "somewhat reliable" for rural and small-town communities in New Hampshire, North Carolina, and Oregon, as well as the USA at large.

The statistical uncertainty of ACS estimates varies considerably between rural and small-town communities and more densely populated urban areas. To illustrate this difference, I compared MHI estimates and the accompanying MOE and CV statistics for three different sized communities: (1) small towns (around 1000 people), (2) urban clusters (around 20,000 people), and (3) urbanized areas (around 100,000 people). As Table 7.5 shows, MHI estimates for urban clusters and urbanized areas are statistically reliable across the states analyzed. Here, CV values range from 1.6 to 7.5. MHI estimates for small towns, on the other hand, are considerably less reliable. Consider, for example, the CV values for two communities in Oregon:

Table 7.4 Coefficient of variation (CV) for median household income (MHI), places with population less than 20,000

		2006–2010	2013–2017
New Hampshire	Mean	20.8	18.1
	Median	18.0	13.7
North Carolina	Mean	22.3	16.7
	Median	14.9	13.6
Oregon	Mean	22.0	16.9
	Median	12.8	12.3
Wisconsin	Mean	13.9	12.5
	Median	9.7	9.6
USA	Mean	22.3	15.8
	Median	13.6	12.4

Source: US Census Bureau (2017 and 2010c), American Community Survey (ACS), 2006–2010 and 2013–2017 (5-year combined estimates). Calculations by author

Table 7.5 Median household income (MHI) estimates with corresponding MOE and CV, 2013–2017

	Estimate	MOE	CV
New Hampshire			
Bethlehem (943)	$53,542	±$10,301	11.7
Portsmouth (21,644)	$72,384	±$4716	4.0
Manchester (110,601)	$56,467	±$1605	1.7
North Carolina			
Aulander (962)	$26,731	±$8033	18.3
Havelock (20,404)	$49,604	±$3012	3.7
High Point (109,849)	$44,642	±$1337	1.8
Oregon			
Drain (931)	$39,583	±$9631	14.8
Roseburg (22,013)	$42,507	±$5258	7.5
Eugene (163,135)	$47,489	±$1357	1.7
Wisconsin			
Amherst (1090)	$45,658	±$7604	10.1
Germantown (19,956)	$79,553	±$6970	5.3
Green Bay (104,796)	$45,473	±$1218	1.6

Source: US Census Bureau (2017), American Community Survey (ACS), 2013–2017 (five-year combined estimates)
Note: MOE values that accompany the population estimates are not reported in Table 7.5. Calculations by author

1.7 for Eugene, OR, and 14.8 for Drain, OR. What are the on-the-ground implications of this statistical difference? In Drain, community officials, planners, and business leaders have to navigate a much larger degree of statistical uncertainty in their MHI estimate. When the MOE is considered, the MHI estimate for Drain ranges from $29,952 ($39,583 – $9631) to $49,214 ($39,583 + $9631). These data underscore the statistical precariousness that many small towns face.

The challenges of statistically uncertain MHI estimates are apparent. Equally important is knowing: (1) to what extent have MHI estimates become more reliable for rural and small towns? and (2) how many rural and small-town communities continue to struggle with unreliable MHI estimates?

First, due to the change in the ACS sampling strategy between the two ACS periods, statistical reliability for rural communities and small towns has improved. Table 7.6 illustrates that MHI estimates have generally become more reliable and the number of communities with unreliable MHI estimates has dropped considerably. In the 2006–2010 period, for example, 11,415 rural and small-town communities had

Table 7.6 CV reliability for median household income (MHI), places <20,000

	2006–2010		2013–2017		Change between periods	
	Number	Percent	Number	Percent	Numeric	Percent
New Hampshire						
Reliable (0–12%)	33	37.5%	38	45.2%	5	15.2%
Somewhat reliable (12–40%)	46	52.3%	37	44.1%	(9)	−19.6%
Unreliable (40%+)	9	10.2%	9	10.7%	0	0.0%
North Carolina						
Reliable (0–12%)	252	36.4%	273	41.1%	21	8.3%
Somewhat reliable (12–40%)	362	52.3%	361	54.3%	(1)	−0.3%
Unreliable (40%+)	78	11.3%	30	4.6%	(48)	−61.5%
Oregon						
Reliable (0–12%)	158	46.9%	153	47.2%	(5)	−3.2%
Somewhat reliable (12–40%)	137	40.6%	143	44.2%	6	4.4%
Unreliable (40%+)	42	12.5%	28	8.6%	(14)	−33.3%
Wisconsin						
Reliable (0–12%)	460	62.7%	464	63.4%	4	0.9%
Somewhat reliable (12–40%)	241	32.8%	247	33.7%	6	2.5%
Unreliable (40%+)	33	4.5%	21	2.9%	(12)	−36.4%
USA						
Reliable (0–12%)	11,415	43.0%	12,068	48.2%	653	5.7%
Somewhat reliable (12–40%)	12,191	46.0%	11,715	46.7%	(476)	−3.9%
Unreliable (40%+)	2923	11.0%	1282	5.1%	(1641)	−56.1%

Source: US Census Bureau (2017 and 2010c), American Community Survey (ACS), 2006–2010 and 2013–2017 (five-year combined estimates). Calculations by author

reliable MHI estimates, compared to 12,068 in the 2013–2017 period. This represents a 5.7% increase in reliable MHI estimates between the two periods. Also, where almost 3,000 rural and small towns had unreliable MHI estimates during the 2006–2010 period, the number dropped by more than half (1282) in the most recent 2013–2017 ACS (Table 7.6). However, the increase in statistical reliability was uneven across the states analyzed. For rural communities and small towns in Oregon and Wisconsin, there was essentially no change in the number of statistically reliable MHI estimates. The largest increases in reliable MHI estimates were among rural and small towns within North Carolina and New Hampshire, increasing by 8.3% and 15.2%, respectively.

Second, despite the improvements in statistical reliability, many rural communities and small towns still suffer from statistically unreliable MHI estimates. More than half (12,997 or 51.8%) of rural and small-town communities across the USA have either somewhat reliable or unreliable MHI estimates in the 2013–2017 ACS (Table 7.6). A greater share of rural and small-town communities in New Hampshire (54.8%), North Carolina (58.9%), and Oregon (52.8%) have somewhat reliable or unreliable MHI estimates. The number of communities across rural America struggling with this issue is significant. Without changes to program eligibility criteria (e.g., considering MOE), changes in sampling strategy, and/or increasing the ACS sample size—pending available government funding—it is essential that states consider policies to assist rural communities and small towns with statistically unreliable MHI estimates.

7.6 Support for Helping Rural Communities with Unreliable ACS Estimates

In Oregon, the state's economic development agency, Business Oregon (through the Infrastructure Finance Authority [IFA]),[11] forges strategic partnerships and makes funding available to rural communities interested in conducting an income census/survey. Oregon IFA, like other state economic development agencies, covers some costs of conducting income censuses/surveys.[12] A February 2015 (p. 7) article in the Oregon League of Cities (LOC) magazine[13] recommends the following approach to unreliable ACS estimates in rural communities in Oregon:

> . . .communities that believe the ACS contains incorrect data and will be adversely impacted should contact their IFA regional coordinator, who can provide guidance and instruction on available options, which may include conducting a local survey. As noted previously, the survey has to be conducted according to the strict requirements of the specific federal agency.

[11]The chief aim of the Oregon IFA is to help Oregon communities apply, receive, and manage federal and state loan/grant funds for water, sewer, roads, and other infrastructure development.

[12]In the recent past, Oregon IFA covered up to $7500 of costs for a census enumeration for cities with a population less than 500 and 50% of costs up to $5000 for a survey with cities with a population of 500 or more. Currently (through June 2019), Oregon IFA covers up to $1000 in costs for conducting a survey.

[13]This link also contains information for conducting an income census/survey (for communities in the Great Lakes RCAP region): http://greatlakesrcap.org/uploads/PDF/Winter2014RCAPConnectionFINAL.pdf

7.7 Summary and Recommendations

State and federal infrastructure grant/loan programs are essential lifelines for improving water, sewerage, and transportation infrastructure for communities across rural America. Eligibility for these programs is often determined according to income estimates from the ACS. The problem, especially for rural communities, is that the statistical uncertainty of ACS data is higher for rural communities, and MOE values are often disregarded for determining program eligibility. This chapter contextualizes the breadth of this problem among small town and rural communities. In the 2013–2017 period, more than half (12,997 or 51.8%) of rural and small-town communities across the USA have either somewhat reliable or unreliable income estimates. Equally important, the reliability of income estimates varies by state. In New Hampshire, North Carolina, and Oregon, for example, a greater share, 54.8%, 58.9%, and 52.8%, respectively, of rural and small-town communities have somewhat reliable or unreliable income estimates.

Communities can hire a consultant to conduct an income census or survey to formally challenge the USCB's ACS estimate. Many rural communities, however, elect not to pursue this approach. Often, they are either unaware that some grant/loan programs reimburse costs for conducting an income census/survey or are dissuaded by the necessary time and resources.

What specific policies can states implement to assist rural communities with unreliable estimates and secure infrastructure funding that advances economically healthy, resilient, and sustainable communities? First, a coordinating agency, presumably the state's economic development agency, should reach out to key partners engaged in rural economic funding development issues (e.g., water, sewerage, transportation, and housing infrastructure) at state, regional, and local levels. Consideration should also be given to partnering with national organizations like the Rural Community Assistance Partnership (RCAP). A 501(c)(3) nonprofit organization, RCAP has contract agreements with the US Departments of Health and Human Services (HHS), USDA Rural Development, and the Environmental Protection Agency (EPA). Another potential national partner is the USCB's State Data Center (SDC) program, which is working collaboratively with the Bureau to assess the viability of Rural Statistical Area (RSA) geographies. This approach minimizes any duplication of efforts by identifying potential partners—both technical and nontechnical—and improving overall efficiency. Second, universities should be considered as census/survey partners. Universities, especially land-grant institutions with rural extension programs or urban-serving institutions working statewide, often have locally specific knowledge from community-engaged scholarship. Having universities involved also provides experiential learning opportunities for students. These learning opportunities may be directly tied to conducting censuses/surveys and may also result from additional community research needs (e.g., housing, transportation, health, and workforce training).

Third, states should adopt guidelines for prioritizing how to best allocate state funds to communities requesting an income census/survey. This is a critically important step because statistical uncertainty of income estimates varies according

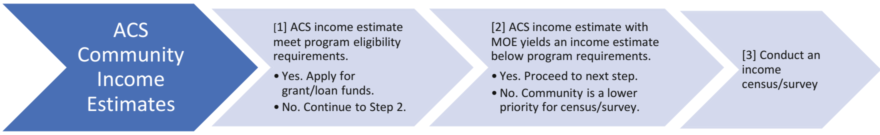


Fig. 7.2 Recommended steps for prioritizing census/survey need

to community size. I recommend states to prioritize community need for an income census/survey according to the approach outlined in Fig. 7.2.

The steps outlined in Fig. 7.2 are as follows:

1. If a community meets program eligibility requirements, then proceed by applying for grant/loan infrastructure funds. If the community does not meet program eligibility requirements, proceed to Step 2.
2. Determine if the ACS income estimate (along with the corresponding MOE) yields an income estimate below the program income threshold.
   A. If a community's ACS income estimate, along with the corresponding MOE, yields an income value below the program income requirement, it means that the "actual" value could qualify the community for grant/loan funding. Proceed to Step 3.
   B. If a community's ACS income estimate, along with the corresponding MOE, *does not* yield an income value below the program income requirement, it means that the "actual" value, more than likely, would *not* qualify the community for grant/loan funding. These communities should be lower priorities for receiving funding for conducting a survey/census.
3. Conduct an income survey/census for communities identified in 2(A).

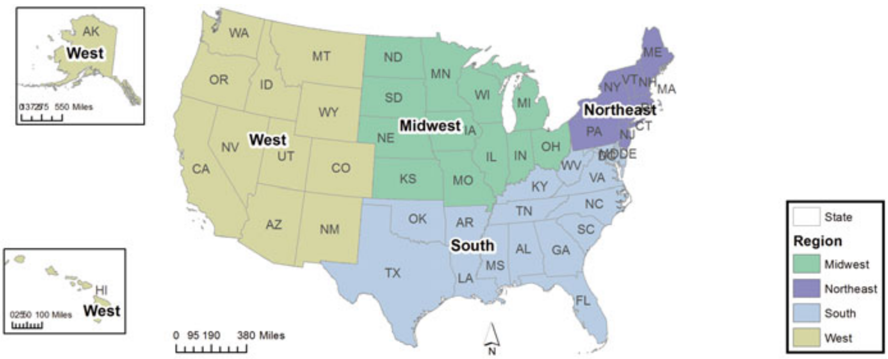
To explain how this proposal would work, consider a grant/loan program with an income estimate threshold of \$35,000. The towns of Admiral and Birdtown each have income estimates of \$36,500, but the MOE for Admiral and Birdtown is ±\$1200 and ±\$3500, respectively. Under this scenario, neither community meets the eligibility requirements for the grant/loan program because MOE is ignored (Step 1).

For Admiral, the income estimate and corresponding MOE *does not* meet the program income requirement (the lower-bound estimate is \$35,300, based on \$36,500 − \$1200) (i.e., Step 2[B]). Also, a test for statistical significance indicates that it is *unlikely* that the "true" income value for Admiral is below the \$35,000 threshold (comparing \$35,000 and \$36,500 [±\$1200] yields a statistically significant t-test value of 2.0563 at 95%). Conversely, Birdtown's income estimate and corresponding MOE *meets* the program income requirement (the lower-bound estimate is \$33,000, based on \$36,500 − \$3500) (i.e., Step 2[A]). And because the estimate for Birdtown contains greater statistical uncertainty (i.e., higher MOE), there is a greater chance–relative to Admiral–that the "true" income value is below the \$35,000 threshold (comparing \$35,000 and \$36,500 (±\$3500) yields a non-statistically significant *t*-test value of 0.7050 at 90, 95, and 99%[14]). In the end, Birdtown is an excellent candidate for receiving funds to conduct an income census/survey. Using these statistical tests can help jurisdictions prioritize the likelihood for whether a community has a "true" income value above/below grant/loan income requirements.

Changes to program eligibility criteria—either explicitly considering MOE or increasing the ACS sample size, for example—would be immediate and effective policy remedies to help rural communities and small towns with unreliable ACS estimates. In reality, however, implementing these changes will take time and resources. States, in the meantime, should invest in policies that efficiently and equitably distribute resources for conducting income surveys/censuses. This approach ensures that rural communities secure much-needed capital projects that advances economic development and quality of life for all citizens and, in particular, for the most vulnerable.

Acknowledgments This research was partially funded with support from Oregon State University (OSU) and Portland State University (PSU). Special thanks to the following individuals, who provided valuable insight and constructive feedback on this manuscript: Paul Lask, Beth Emshoff, Nick Chun, Lena Etuk, Mallory Rahe, Charles Rynerson, Jeff Sherman, David Tetrick, Kevin Tracy, and Bruce Weber. Thanks also to Business Oregon. All errors and oversights are solely the responsibility of the author.

Appendix: Census Regions



Source: US Census Bureau

[14]This example uses 90% as the cutoff for statistical significance based on prevailing practice in social science research, but determining the level of statistical significance (i.e., how much uncertainty one is willing to tolerate) is somewhat arbitrary, and ultimately up to each state.

References

Alvarez JA, Salvo J (2014) Navigating reliability of small area data, lessons from New York City. Presented at American Community Survey (ACS) data user group conference, May 29–30, Washington, DC

Asiala, ME (2012) American Community Survey sampling and weighting. Presented at measuring people in place conference, Colorado University, Boulder. 5 October 2012. https://cupc.colorado.edu/workshops/measuring_people_in_place/themes/theme1/asiala.pdf

Carpenter R, Reamer A (2010) Surveying for dollars: the role of the American Community Survey in the geographic distribution of federal funds. Brookings Institution, Washington, DC. https://www.brookings.edu/research/surveying-for-dollars-the-role-of-the-american-community-survey-in-the-geographic-distribution-of-federal-funds/

Caster P (2017) House committee passes Herrera-Beutler bill on flawed income data. http://www.chronline.com/news/house-committee-passes-herrera-beutler-bill-on-flawed-income-data/article_e8c3a330-6d6f-11e7-848b-8f4c80959193.html. Accessed 14 Feb 2019

Citro C, Kalton G (eds) (2007) Using the American Community Survey: benefits and challenges. National Academies Press, Washington, DC. http://www.nap.edu/read/11901/chapter/1#ii

Cromartie J, Swanson L (1996) Defining metropolitan areas and the urban-rural continuum: a comparison of statistical areas based on county and sub-county geography. Staff paper no. AGES9603. Economic Research Service (ERS), U.S. Department of Agriculture, Washington DC

Environmental Systems Research Institute (ESRI) (2014) The American Community Survey. https://www.esri.com/library/whitepapers/pdfs/the-american-community-survey.pdf. Accessed 21 Nov 2018

Heuvelink GBM, Burrough PA (2002) Developments in statistical approaches to spatial uncertainty and its propagation. Int J Geogr Inf Sci 16(2):111–113

Isserman AM (2005) In the national interest: defining rural and urban correctly in research and public policy. Int Reg Sci Rev 28(4):465–499

Jurjevich JR, Griffin AG, Spielman SE, Folch DC, Merrick M, Nagle NN (2018) Navigating statistical uncertainty: how urban and regional planners understand and work with American Community Survey (ACS) data for guiding policy. J Am Plan Assoc 84(2):112–126

McGranahan DA, Hession JC, Hines FK, Jordan MF (1986) Social and economic characteristics of the population in metro and nonmetro counties. 1 970–80 RDR-58. Economic Research Service. U.S. Department of Agriculture, Washington, DC

National Academy of Sciences (2015) Realizing the potential of the American Community Survey: challenges, tradeoffs, and opportunities. National Academies Press, Washington DC

Navarro F (2012) An introduction to ACS statistical methods and lessons learned. Presented at measuring people in place conference. Colorado University, Boulder. 5 October 2012. http://www.colorado.edu/ibs/cupc/workshops/measuring_people_in_place/themes/theme1/navarro.pdf

Nesse K, Rahe M (2015) Conflicts in the use of ACS by federal agencies between statutory requirements and survey methodology. Popul Res Policy Rev 34(4):361–480. https://doi.org/10.1007/s11113-014-9352-7

Oregon League of Cities (LOC) (2015) IFA Offers Guidance on American Community Survey. Available at: http://www.orcities.org/Portals/17/Publications/localfocus/Feb2015web.pdf. Accessed 4 Jan 2019

Ratcliffe M, Burd C, Holder K, Fields A (2016) Defining rural at the U.S. Census Bureau. ACSGEO-1. U.S. Census Bureau, Washington, DC

Spielman SE, Folch DC (2015) Reducing uncertainty in the American Community Survey through data-driven regionalization. PLoS One 10:e0115626. https://doi.org/10.1371/journal.pone.0115626

Spielman SE, Folch DC, Nagle NN (2014) Patterns and causes of uncertainty in the American Community Survey (ACS). Appl Geogr 46:47–157

Torrieri NK (2007) America is changing, and so is the census. Am Stat 61(1):16–21
U.S. Census Bureau (2009) A compass for understanding and using American Community Survey data: what state and local governments need to know. US Government Printing Office, Washington, DC
U.S. Census Bureau (2010a) Urban and rural (P2): Census 2010, summary file 1 (SF1), 100 percent data
U.S. Census Bureau (2010b) Rural-urban classification criteria. Available at: https://www.census.gov/geo/reference/ua/urban-rural-2010.html
U.S. Census Bureau (2010c) American Community Survey (ACS), 2006–2010 combined 5-Year estimates, B19013: median household income (dollars). Generated using American FactFinder. Accessed 14 Sept 2012
U.S. Census Bureau (2017) American Community Survey (ACS), 2013–2017 combined 5-Year estimates, B19013: median household income (dollars). Generated using American FactFinder. Accessed 3 Mar 2019)
U.S. Department of Agriculture (2012) Administrative Notice (AN) 4679. Implementation of 2010 Census Data for Rural Development Programs. http://www.ruralhome.org/storage/documents/rd_obligations/rd_docs/an4679.pdf. Accessed 30 Nov 2019
US Census Bureau (2011) American Community Survey: accuracy of the data. https://www2.census.gov/programs-surveys/acs/tech_docs/accuracy/ACS_Accuracy_of_Data_2011.pdf. Accessed 5 Apr 2019
US Census Bureau (2019) History: questionnaires. https://www.census.gov/history/www/through_the_decades/questionnaires/. Accessed 3 Mar 2019
Usowski K, Rodgers L, Lihn M, Kahn P (2008) Using American Community Survey data in HUD's income limits and fair market rents. Cityscape J Policy Dev Res 10(1):199–210

Chapter 8
Unpacking the Nature of Long-Term Residential Stability

William A. V. Clark and William Lisowski

Abstract This study is motivated by the increasing calls in the literature to understand long-term residential staying. We take up this issue using longitudinal data from the Household, Income and Labour Dynamics in Australia (HILDA) survey. Until recently, the focus in life course studies has been on events rather than states and, in the case of mobility, on moving rather than the duration of stay between moves. Because most people in fact move quite infrequently, it is natural to ask—can we identify what lies beneath the decision *not* to move? The value of turning our attention to the decision to stay is that it refocuses attention on the people/place connection—on people staying in places and the connections they have to their local environments. We show that staying is embedded in family stability and is a direct response to long-term family and household stability. Clearly, aging matters—we are less likely to move as we age—but the research shows that both family structure and place play roles in staying, with varying outcomes for couples and singles. As previous research has also documented, the presence of children plays an important role in staying as does the stability of ownership.

Keywords Duration · Family stability · Stayers · Life cycle

8.1 Introduction

There are increasing calls to expand the focus from residential moves to studies of the decision to stay and residential duration more generally. Of course we do not want to go back to a simple mover/stayer dichotomous discussion; rather we need to think about the continuum of behaviors which include both staying and moving.

W. A. V. Clark (✉)
University of California, Los Angeles, CA, USA
e-mail: wclark@geog.ucla.edu

W. Lisowski
Easton, PA, USA

R. S. Franklin (ed.), *Population, Place, and Spatial Interaction*, New Frontiers in Regional Science: Asian Perspectives 40,
https://doi.org/10.1007/978-981-13-9231-3_8

As we progress through the life course, it is possible to think of staying in a familiar place as the long-term state which has as "book ends" changes in status, which can involve a change in family composition, a change in location, or both. In this structure moving and staying are part of a process of change over the life course.

Mobility has been the focus of most research efforts, and much less attention has been paid to the time between mobility events. That is, the focus has been primarily on the event of moving rather than the spell of staying. In this analysis we explore the staying process within the context of the life course and bring more understanding to what underlies long-term staying and, when those households do move after lengthy stays, what helps us understand the process. We have a well-documented understanding of what generates moves but a less well-developed understanding of what happens in between the moves. So, in this discussion we ask who stays for long periods between moves, and what are the explanations for long-term staying? Of course mobility is part of the process we are examining, but it is not mobility and events that are the focus; rather we are trying to do something different in capturing the nature of long-term staying and its characteristics. We emphasize that this is not simply the reverse of the focus on the event (usually a 1-year focus, or pooled 1-year events); it is a study of the *interval* of staying, in our case a 12-year interval of staying, in the study of families in Australia. We are turning the attention from the often 1-year event of moving or not to a varying multi-year spell of staying.

By virtue of their predominance in the population, it is stayers who exercise the major influence on neighborhood social structure simply by aging in place. As Coulter (2013) and Coulter et al. (2015) note, by redirecting attention to stayers, we are focusing not on a single point in time but on "relational practices" where it is not discrete events which are important but the nature of linked lives across space and time, in this case the time of staying. In this view "spatial moorings," rather than the event of moving, redirect our attention to the locational anchors around which we organize our daily lives. Of course all this is only possible with longitudinal data sets, which allow us to view and analyze intervals of staying. It directs attention to the complex connections between family dynamics and housing and locality satisfaction.

Duration (in the residence or the neighborhood) has always mattered for a whole range of outcomes from stability in schooling to the quality of social interactions. And duration has taken on a new immediacy with declining mobility rates and the changing nature of "rootedness" (Cooke 2013). It is also a focus in our rapidly aging societies, where the tendency to stay is growing along with aging (Fernández-Carro and Evandrou 2014). Theoretically, the focus on duration has been embedded in the "endowment effect"—the idea that households become attached to their locations, which means they build up an endowment and are therefore averse to moving without significant compensation (Morrison and Clark 2016; Clark and Lisowski 2017). The role of attachment is further emphasized by recent studies in which surveys report that Americans want to stay in their own homes, close to their families and near friends. Nearly two-thirds of baby boomers have no plans to move at all. They will "age in place" in homes and communities where they have often lived for a decade or more (Osterland 2015). Even those who move are staying in or near their metro area (Burbank and Keely 2015).

But while we have a growing sense that duration in dwellings and in neighborhoods matters, and is increasing, we know less about how the decisions to stay fit into the broader context of the life course. Thus, we use data on a large pool of individuals who have multiple records in a 13-year period and then focus in on the behavior of individuals in households. We can also answer the question just how much mobility is created by one set of life course events, those of marriage, cohabitation, divorce, separation, and widowhood which almost all result in residential moves. Then, by focusing on a cohort without formations or dissolutions, we can narrow our perspective to evaluate how long-term stayers fit into the larger life course experiential structure. The end result is some of the first evidence of the proportional effects of family formation and dissolution on the mobility process and for those continuing households what motivates their staying or moving behavior.

8.2 The Theoretical Context and Previous Research

Much of the research on mobility, even those studies which use longitudinal data, in essence study only the event (as we noted earlier, usually pooled 1-year events) and do not examine the continuum of spells over, at a minimum, several years. Theoretically, however, it is these long periods with multiple moves that are at the heart of understanding how mobility and immobility intersect. In the analysis in this chapter, there are two important dimensions which underlie the analysis of events and spells. We are interested in the proportional role of family composition in the process of staying and moving and in who stays for long periods and who for shorter periods.

Not staying has always been dominated by the young. Young renters account for much of the totality of movement. A wide range of studies emphasize the high level of movement between rental units in the early household formation stage and how it creates much of the overall mobility in a population. A whole series of studies of family panels showed that the young are mobile and that a household's propensity to change place of residence is a function of the need to adjust housing consumption, and this often occurred in the transition from being a renter to being an owner (Ioannides 1987; Clark and Dieleman 1996; Mulder 1993).

Family structure and family connections, too, have been at the heart of studies of the likelihood of staying. There are changing linkages between children and older parents which in turn will likely increase the need for local connections and by extension on staying in local areas if not in the same residence (Michelin et al. 2008; Mulder and Malmberg 2014). Because decisions about moving or staying are increasingly embedded in the larger context of extended family structures, now households, adapting to the extended life spans of older parents, have to make decisions about the needs of elderly parents, the desires of grandparents, and the whole set of familial connections (Taylor et al. 2008; Cohn and Morin 2008). A Pew survey finds that stayers overwhelmingly say they remain because of family ties and because their hometowns are good places to raise children. Their life circumstances match those explanations. Most stayers say at least half a dozen members of their extended families live within an hour's drive; for 40%, more than 10 relatives live nearby (Cohn and Morin 2008). Stayers also cite a feeling of belonging as a major reason for staying put.

Two changes in work–residence links may also be playing a role in changing duration. First, the increase in two-worker households makes it more difficult to uproot and move (Cooke 2013). We know that today's two-earner married households are about 46% less likely to move across state lines than were their counterparts in the 1980s (Taylor et al. 2008). Second, a growing increase in home-based work for at least a segment of the employed population and the associated ability to use complex technology without regularly attending a specific location has made it possible "not to have to move" and further increased the likelihood of staying locally.

All of this directs us to the questions about why and how people stay and to how we might better understand this important process. To do this, an initial study created a formal model of staying by focusing on how moving and transaction costs tend to discourage moving especially if you are a homeowner (Goodman 2003). From the Goodman perspective, it takes more than a small change in income to generate a move, other things being equal, and thus staying is more likely on the whole than moving (Goodman 2002). Goodman echoed earlier work by Morrison (1967, 1970) on duration of stay, when he concluded that "length of stay has measurable and important effects" on the probability of future staying (Goodman 2003, p. 107).

Despite these early studies, there has been remarkably little sustained quantitative work devoted to understanding staying. The studies of cumulative inertia, while useful, did not unpack the underlying dimensions of staying (Clark and Huff 1977). However, staying and duration in the same occupancy has been important in the studies of attachment, both attachment to the individual dwelling and attachment to the neighborhood (Clark et al. 2015). These studies emphasize the attachment links of household members to their neighborhood—their familiarity with the area, their social ties, and their feelings of security—all of which increase with length of residence. It is a sense of community which plays an important role in their decisions and the studies that explore staying in any depth point to the role of "roots" and spatial bonds in the neighborhood, connectedness to family and friends in the local area, and feelings of security and overall well-being (Uzzell et al. 2002; Woldoff 2002; Lewicka 2011). Fischer and Malmberg (2001) also draw attention to the role of place and local attachment as explanation for staying as does Hanson (2005). People with ties are less prone to move, and they point to what they call "location-specific insider advantages" in the decision to stay. Over longer periods, advantages accrue to stayers who accumulate advantages that are nontransferable, that is, they increase their endowment (Clark and Lisowski 2017).

The turn to a focus on staying has been especially focused on the very long-term stayers and studies of elderly staying in particular. These studies often couched in terms of "sedentism" do provide another window on the role of roots, both in terms of family and links to the locality (Hjalm 2014). A study along similar lines of very long-term stayers across Europe also noted the importance of roots and homeownership, a finding that we will explore in our analysis. They also reiterated that staying is not straightforward and that older stayers can be both intentional and unintentional stayers, and for the latter, clearly attachment is not the motivating variable (Fernández-Carro and Evandrou 2014).

To begin an analysis of these questions about staying, we position our analysis within the context of stable family structures over the life course. We go beyond the usual focus on family disruption as a motivating force for mobility to how connections to dwellings and the neighborhood forge the likelihood of staying. We specifically structure the sample to report the variations in staying across different populations and show how stable compositions in family structures are a central element of the staying process. To reiterate our focus, we hypothesize that for stable families, staying is strongly related to (a) the presence of children, (b) ownership, and (c) high levels of satisfaction with the dwelling or neighborhood.

8.3 Life Cycle Changes and Duration

The data to examine socio-spatial mobility in Australia come from the first 13 waves (2001–2013) of the Household, Income and Labour Dynamics in Australia (HILDA) survey, a longitudinal survey of approximately 7400 households with approximately 19900 respondents. The survey is modeled on and is similar to the Panel Study of Income Dynamics (PSID) in the USA and the British Household Panel Survey (BHPS, now called the "Understanding Society" study). The HILDA survey has detailed data on household composition, economic characteristics of households, mobility and migration, and a wide range of subjective measures of place and data on family change. The source of our data is HILDA Release 13 from November 2014 (Summerfield et al. 2014), containing data from the first 13 waves.[1]

Families are identified from the original survey using the individual and partner cross-wave IDs and come into and out of existence throughout HILDA's 13 waves. We identified a subset of 4003 households (families or individuals) from the HILDA survey who had data across 11 to 13 waves of the HILDA data, including wave 1 and wave 13. We used 11 plus waves to increase sample size and where we could interpolate data. Of these, 3145 are our couples and singles who did not have a family change (what we will call our stable couples and singles), plus 858 who experienced a family composition change over the 11–13 waves. Of the original sample those in wave 1, 3285 were not in wave 13 mainly due to family dissolution and drop out. We are reasonably certain that 2987 with less than 11 waves were lost due to dissolutions and moves.

[1] Attrition in longitudinal surveys is always an issue, and while the rates from the HILDA survey are somewhat similar to other surveys like the BHPS, there is a higher attrition of younger single households which will modestly affect the outcomes (see Watson and Wooden 2004 for a discussion).

Table 8.1 Staying and moving by stable families in the HILDA sample

	All	Stayers		Movers	
Stable couples	2139	1130	52.8%	1009	47.2%
Stable singles	1006	519	51.6%	487	48.4%
All stable families	3145	1649	52.4%	1496	47.6%

Table 8.2 Duration of stable households at wave 13

	Owner		Renter		Total	
Duration (years)	N	Percent	N	Percent	N	Percent
0–3	359	15.2	236	31.7	595	19.1
4–8	340	14.4	207	27.8	547	17.6
9–14	470	19.8	166	22.3	636	20.4
15–20	406	17.1	84	11.3	490	15.7
21+	793	33.5	51	6.9	844	27.1
Total	2368		744		3112	

The calculation of duration included reported data on previous residence location established at Wave 1

The takeaway from this analysis is just how much dynamism there is in family structures and how this is linked to residential change. In this longitudinal data set, about half of the sample was lost to family change and moves. If we add the 858 families for which we have continuing data but who had composition change to the 2987 households with a probable dissolution, we can reiterate that over quite long periods much of the mobility in the system is due directly to family composition change. Our standard models of mobility show how divorce, separation, and marriage are significant explanatory variables in models of mobility but what this analysis adds is a measure of the numerical dimension of status change.

Households with stable composition also move as we show in Table 8.1. But, how long do they stay on average between moves? Table 8.2 provides basic data for owners and renters and their durations. Stable families move as we demonstrate in Table 8.1, but stable families are likely to have long durations (Table 8.2).[2] More than half of owners have durations of more than 15 years, and 70% have 9 plus years in their residence. The distribution for renters who are in stable families is lower overall as we would expect from the greater general transiency of renters. Still, 40% have been in the dwelling for 9 plus years, and nearly a fifth of renters in stable families have durations greater than 15 years. Longitudinal data provide a very different perspective on staying and on the differences between owners and renters. It is not that the cohort of stable families always has long durations, in fact 15% of owners and 20% of renters have durations in the window of 0–3 years, and in the next section, we further explore the differences between stable families and those with family compositional change.

[2]Table 8.2 excludes a very small fraction of cases for which duration was not known at that time being measured—this amounts to just 33 of 3145 families.

8.4 Explaining Duration

8.4.1 Descriptive Interpretations of Staying

We first document the variation of staying across the three measures of demographic characteristics, status and economic position, and satisfaction and place attachment. We do this separately for stable couples and stable singles. The three classes of variables are the variables which will be used in the next section to model the probability of staying.

Age and family structure are important dimensions of stability. For both couples and singles, age, as we know from much research on mobility, is powerfully associated with the decision to stay or to move. Of those under 35 in wave 1, only 24% for couples stayed during the window and less than 20% in the case of singles. At the other extreme is a very high probability of staying for older couples and older singles. Within the demographic variables, it is the role of children that dramatically differentiates staying and moving for both couples and singles. Couples who had a child at the beginning of the survey interval have high probabilities of staying, but it was just the opposite for singles. For couples as well as single parents, households that gain a child have a much higher likelihood of moving. While there is overall a tendency to stay for couples with a child, when a child or an additional child enters, we can see the effect of having a child and the classic interpretation of needing more space in the dwelling.

Within the status measures, it is housing tenure that has the most powerful impact on staying or moving, but job change also plays a role. Stayers are owners by and large, and we see the powerful role in terms of the likelihood of staying over long periods when we recall that the stayers are those in this case who have not moved over the 13 waves in the sample. There are not particularly big differences across the occupations although clerical workers who were couples are more likely to stay as are those who are not employed, and the results are similar for singles. Members of couples in clerical occupations may be members of dual-income households, and likely also to be owners and hence also likely to be stayers. Education has only a modest role to play in staying or moving. Those couples with a BA were several percentage points more likely to stay than those with a high school education, and for singles the education affect seems to be more for those with less than high school (Table 8.3).

Income (measured in quintiles) as expected creates greater opportunities to move, and consequently the higher income quintile has lower rates of staying. For singles those in lower and higher income quintiles are likely to stay but again by only a modestly higher rate. Income change plays some role mainly for those with a very significant income gain that of course makes moving possible and opens up a wide range of alternatives to those of staying in the house or neighborhood. While tenure equals staying, job change equals moving and confirms work from other studies of the link between mobility and job change (Clark and Davies-Withers 1999).

Table 8.3 Demographics of stable couples and stable singles with 11+ waves in the HILDA survey

	Stable couples			Stable singles		
	All	Stayers	%	All	Stayers	%
Total population	2139	1130	52.8	1006	519	51.6
Average number of moves for non-stayers		1.8			2.2	
Age						
<35	356	86	24.2	112	22	19.6
35–45	601	312	51.9	220	99	45.0
45–55	559	334	59.7	239	117	49.0
55–65	425	256	60.2	216	136	63.0
65+	198	142	71.7	219	145	66.2
Single individual gender						
Male				305	166	54.4
Female				701	353	50.4
Presence of children						
None	697	397	57.0	692	381	55.1
Wave 1 (beginning)	423	252	59.6	153	61	39.9
Wave 1 and 13 (continuing)	827	434	52.5	139	68	48.9
Wave 13 (child added)	192	47	24.5	22	9	40.9
Highest education						
Less than HS	746	359	48.1	165	77	46.7
High school graduate	267	152	56.9	92	53	57.6
Diploma	765	406	53.1	267	144	53.9
BA	361	213	59.0	482	245	50.8
Occupation						
Professional	1056	522	49.4	187	95	50.8
Technical	434	215	49.5	125	58	46.4
Clerical	240	127	52.9	147	56	38.1
Laborers	161	99	61.5	149	62	41.6
Not employed	248	167	67.3	398	248	62.3
Household income						
Lowest quintile	145	86	59.3	404	219	54.2
Second quintile	291	158	54.3	259	121	46.7
Middle quintile	394	209	53.0	206	102	49.5
Fourth quintile	627	315	50.2	102	59	57.8
Highest quintile	682	362	53.1	35	18	51.4
Household income change						
Lowest quintile	428	240	56.1	202	103	51.0
Second quintile	428	250	58.4	201	100	49.8
Middle quintile	428	246	57.5	201	117	58.2
Fourth quintile	428	205	47.9	201	107	53.2
Highest quintile	427	189	44.3	201	92	45.8
Neighborhood advantage/disadvantage						
Lowest quintile	382	193	50.5	269	141	52.4
Second quintile	419	214	51.1	232	111	47.8

(continued)

Table 8.3 (continued)

	Stable couples			Stable singles		
	All	Stayers	%	All	Stayers	%
Middle quintile	425	229	53.9	150	80	53.3
Fourth quintile	426	239	56.1	159	87	54.7
Highest quintile	487	255	52.4	196	100	51.0
Housing tenure						
Owner	1779	1068	60.0	612	384	62.7
Renter	360	62	17.2	394	135	34.3
Employment history						
Single employer	782	518	66.2	623	381	61.2
Multiple employers	1357	612	45.1	383	138	36.0

Table 8.4 Place and dwelling attachment

	Stable couples			Stable singles		
	All	Stayers	%	All	Stayers	%
Housing satisfaction						
Lowest quintile	428	183	42.8	210	86	41.0
Second quintile	428	200	46.7	215	98	45.6
Middle quintile	430	207	48.1	190	99	52.1
Fourth quintile	433	269	62.1	190	103	54.2
Highest quintile	420	271	64.5	201	133	66.2
Neighborhood satisfaction						
Lowest quintile	437	202	46.2	211	90	42.7
Second quintile	422	198	46.9	209	96	45.9
Middle quintile	437	240	54.9	209	93	44.5
Fourth quintile	427	231	54.1	192	119	62.0
Highest quintile	416	259	62.3	185	121	65.4

Finally, we examine the impact of satisfaction and place attachment (Table 8.4). Where the couple or a single person lived and the levels of satisfaction with the dwelling and the neighborhood do have associations with staying. High levels of satisfaction are associated with high proportions staying, and they increase monotonically from those with low levels of satisfaction. The table provides a nice explication of the role of satisfaction in staying. The findings are replicated for both the dwelling and the neighborhood and for couples and singles. The difference between those who have low levels of satisfaction with the dwelling and neighborhood and those who report high levels of dwelling and place attachment is more than 20 points. Clearly, dwellings and places matter, and of course they are interrelated.

However, it is not a uniform outcome, and many do move despite high levels of satisfaction with both house and neighborhood—nearly a third of those with high levels of satisfaction still make residential changes. This is an expected finding because you can be satisfied on environmental levels, for example, but still need more space, and in turn it lowers the explanatory power of the satisfaction variables. Those with low levels of satisfaction are much less likely to stay suggesting that it is a push factor rather than a holding factor.

8.4.2 Multivariate Analysis of Staying

In our analytical universe of stable families present in at least 11 waves, including both waves 1 and 13, the dependent variable is whether the family did not make a residential move between waves 1 and 13. The variables that are predictors of the likelihood of staying include demographic measures of the household (age, education, and children's presence and role), variables that are measures of status and economic position (income, housing tenure, occupation, and job change) and measures of location and satisfaction with the dwelling and the neighborhood. We exclude households where there are family compositional changes and thus do not measure the triggering events like marriage or cohabitation and divorce or separation as they uniformly lead to residential change.[3]

The independent variables that measure household and family characteristics include age, sex (for single individuals), family, marital, socioeconomic (income and education), and employment status. The housing tenure variable captures the important issue of housing stability, and a set of measures on satisfaction are included to measure the link and satisfaction with locality. The usual pooled cross-sectional analysis tends to measure causes and effects in close proximity of time, while the decision to move is often taken over time and perhaps even in advance of the triggering event (e.g., to prepare for anticipated children). We instead create variables that summarize the family's characteristics *over the 13 waves of survey data*. Doing so presents two sets of challenges: summarizing values from multiple waves and combining values for two members of a couple. Some variables are relatively straightforward. For example, we measure age in wave 1, and for a stable couple, the older of the two ages is used. Similarly, we measure household income in wave 1, which is the same for both members of a couple. For household income change over the 13 waves, we calculate the difference between wave 13 income and wave 1 income, adjusting to constant dollars. Other variables are less straightforward. We measure whether job change since the previous interview has occurred in any of waves 2–13 and for couples, for either member. For measures of satisfaction

[3]Detailed notes on the construction of the variables are reported in a Table 8.7 in the Appendix.

with housing and neighborhood, we average the values (for both members of a couple) over the 13 waves of data.[4] While these measures can occur at any point in time, relative to a move, our focus is not on causality but on whether a job change at some point in the interval of observation, or overall long duration dissatisfaction with housing and neighborhood, is associated with a change in location.

The models of staying are first presented as a sequence in which first demographic variables are entered followed by status measures and satisfaction measures. This approach allows us to calculate the additive impact of adding status and satisfaction measures after the demographic explanations have been entered (Table 8.5). Tests of collective significance are presented for variables (e.g., age) rather than separate tests for individual measures (e.g., for age and for age squared). A second table (Table 8.6) provides the log odds ratios for the individual measures in the full models.

As expected age and age squared are powerful predictors of the probability of staying or moving and in combination with the presence of children provide the context within which staying takes place. Education is not significant. When we add status measures to the demographic measures, both housing tenure and job change play important roles in predicting whether a family stays or moves. Ownership and job change are clearly the most important predictors of the likelihood of staying as measured by the log odds in the individual variables (Table 8.6), with job change reducing the likelihood of staying. For couples, both income change and satisfaction are significant at the.10 level which confirms that satisfaction does matter, but the additional explanation from including satisfaction is modest. That said, the ownership variable is likely to be masking any direct effect of satisfaction of the dwelling. For singles, neighborhood is significant and adds to the overall explanation of the models. In the multivariate model, the strong descriptive findings of place effects are more muted. Place effects play out at the margin, while the measures of family status and ownership capture the decision to stay or move.

Staying is a demographic process linked closely to ownership, and in this sense this chapter has reiterated findings of the role of tenure, but it has added important contributions on how children play a role in the likelihood of staying. Viewing the table of log odds for the individual measures on children, we find that couples who have children at the initial wave have relatively high probabilities of staying (Table 8.6). We can infer that they had made decisions about living space and the birth of children in the window before the first wave. In both the initial and ending waves where there are children, there are very high log odds of staying. Children were a continuing part of the process of staying. The continuing presence of children then is a powerful predictor of staying as is demonstrated in this analysis using

[4]Averaging would be a problem if we were looking at causation, but we are looking at long-run satisfaction with housing and neighborhood in relationship to moving or staying. Additionally, 90 % of the individuals in established families who did not move the association of year to year satisfaction is high.

Table 8.5 Explanatory models of staying

	d.f.	χ^2	Signif.	χ^2	Signif.	χ^2	Signif.
Stable couples							
Demography							
Age and age squared	2	132.93	0.000***	50.76	0.000***	44.28	0.000***
Presence of children	3	51.13	0.000***	40.94	0.000***	40.33	0.000***
Highest education	3	5.39	0.145	3.87	0.276	3.06	0.382
Status							
Occupation	4			6.00	0.199	6.01	0.198
Household income	4			2.11	0.716	2.22	0.695
Household income change	4			9.07	0.059	8.36	0.079
Housing tenure	1			120.01	0.000***	116.60	0.000***
Job change	1			19.27	0.000***	18.55	0.000***
Advantage/disadvantage	4			3.06	0.547	2.71	0.607
Satisfaction							
Housing satisfaction	4					6.50	0.165
Neighborhood satisfaction	4					4.02	0.403
Satisfaction measures	8					13.77	0.088
Model summary							
Likelihood ratio χ^2		235		424		438	
Degrees of freedom		8		26		34	
Prob > χ^2		0.000		0.000		0.000	
Pseudo R^2		0.079		0.144		0.148	
Number of observations		2139		2139		2139	
Stable singles							
Demography							
Age and age squared	2	79.10	0.000***	27.43	0.000***	23.48	0.000***
Single individual gender	1	2.44	0.118	2.55	0.110	3.24	0.072

Presence of children	3	14.38	0.002**	12.48	0.006**	13.02	0.005**
Highest education	3	5.04	0.169	3.96	0.266	3.76	0.288
Status							
Occupation	4			4.80	0.309	3.98	0.408
Household income	4			1.54	0.820	2.20	0.699
Household income change	4			1.30	0.862	1.32	0.859
Housing tenure	1			33.45	0.000***	31.03	0.000***
Job change	1			5.83	0.016*	5.01	0.025*
Advantage/disadvantage	4			1.13	0.890	1.31	0.860
Satisfaction							
Housing satisfaction	4					1.12	0.891
Neighborhood satisfaction	4					12.67	0.013*
Satisfaction measures	8					17.71	0.023*
Model summary							
Likelihood ratio χ^2		112		170		188	
Degrees of freedom		9		27		35	
Prob > χ^2		0.000		0.000		0.000	
Pseudo R^2		0.080		0.122		0.135	
Number of observations		1006		1006		1006	

Notes: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 8.6 Logistic regression results for the full models of staying, presented as odds ratios

	Stable couples			Stable singles		
	Odds ratio	Std. Err.	Z	Odds ratio	Std. Err.	Z
Age						
Age	1.131	0.040	3.44***	1.172	0.043	4.37***
Age squared	0.999	0.000	−2.21*	0.999	0.000	−3.85***
Single individual gender (ref male)						
Female				0.730	0.128	−1.80*
Presence of children (ref neither waves 1 nor 13)						
In wave 13	1.551	0.420	1.62	0.647	0.327	−0.86
In wave 1	1.747	0.268	3.63***	0.680	0.153	−1.71*
In waves 1 and 13	2.757	0.471	5.93***	1.868	0.483	2.42*
Highest education (ref less than HS)						
BA or more	0.789	0.137	−1.36	0.834	0.204	−0.74
Diploma	1.017	0.193	0.09	1.158	0.309	0.55
HS grad or cert.	0.913	0.138	−0.61	1.286	0.235	1.38
Occupation (ref laborers/operators)						
Professional	0.628	0.138	−2.12*	1.287	0.371	0.87
Technical/service	0.607	0.135	−2.24*	1.201	0.337	0.65
Clerical/sales	0.657	0.158	−1.75*	0.831	0.229	−0.67
Not employed	0.581	0.157	−2.01*	1.305	0.366	0.95
Household income (ref middle quintile)						
Lowest quintile	0.924	0.228	−0.32	0.796	0.196	−0.93
Second quintile	0.780	0.148	−1.31	0.918	0.204	−0.39
Fourth quintile	0.846	0.127	−1.11	1.177	0.324	0.59
Highest quintile	0.884	0.145	−0.75	0.754	0.312	−0.68
Household income change (ref middle quintile)						
Lowest quintile	0.764	0.135	−1.52	0.779	0.209	−0.93
Second quintile	1.028	0.166	0.17	0.897	0.213	−0.46
Fourth quintile	0.843	0.136	−1.06	0.965	0.220	−0.16
Highest quintile	0.685	0.115	−2.26*	0.807	0.204	−0.85
Housing tenure (ref renter)						
Owner	5.758	0.933	10.80***	2.425	0.386	5.57***
Job change (ref single or no employer)						
2 or more empl.	0.597	0.071	−4.31***	0.647	0.126	−2.24*
Advantage/disadvantage (ref middle quintile)						
Lowest quintile	0.895	0.145	−0.69	1.032	0.237	0.14
Second quintile	1.012	0.157	0.08	0.930	0.219	−0.31
Fourth quintile	1.160	0.181	0.95	1.211	0.308	0.75
Highest quintile	0.971	0.153	−0.19	1.055	0.263	0.21

(continued)

Table 8.6 (continued)

	Stable couples			Stable singles		
	Odds ratio	Std. Err.	Z	Odds ratio	Std. Err.	Z
Housing satisfaction (ref middle quintile)						
Lowest quintile	1.224	0.200	1.24	1.012	0.248	0.05
Second quintile	1.115	0.170	0.71	0.943	0.212	−0.26
Fourth quintile	1.438	0.222	2.35*	0.825	0.189	−0.84
Highest quintile	1.349	0.227	1.78*	1.004	0.254	0.01
Neighborhood satisfaction (ref middle quintile)						
Lowest quintile	0.837	0.137	−1.09	0.992	0.232	−0.04
Second quintile	0.774	0.120	−1.65*	1.042	0.231	0.19
Fourth quintile	0.834	0.129	−1.18	1.893	0.436	2.77**
Highest quintile	0.994	0.163	−0.03	1.834	0.461	2.41*
Model summary						
Likelihood ratio χ^2	438			188		
Degrees of freedom	34			35		
Prob > χ^2	0.000			0.000		
Pseudo R^2	0.148			0.135		
Observations	2139			1006		

Notes: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

longitudinal data rather than cross-sectional analysis. This is a new finding of just how children play a role in the likelihood of staying.

For singles the results are similar with respect to the role of children. However, the overall levels of prediction are slightly lower. It is also useful to specifically focus on the role of children for singles. Where children are present in waves 1 and 13, the odds are significantly greater that the single household will not move, but the entry of a child between wave 1 and wave 13 has a significantly negative effect on the likelihood of staying. Clearly this is an adaptive process to restructuring households.

8.5 Observations and Conclusions

This research reported here provides a window on the process of long-term staying and sets it within the relational framework of moving and staying. The models explore the associations with who stays and the relative role of demographic characteristics, status, and satisfaction in the process of staying. The study breaks new ground in that we pursue behavior over a longer window that is common in studies of the binary event of either moving or staying, and it is also innovative by linking family stability to behavior. The study moves away from either logit models of the event of move/stay or models which focus on the length of duration. Rather we focus on duration over a decade-long period and distinguish that cohort of stayers from those who interrupt their locations with one or more frequent moves.

The study is important in the sense that it is not simply the reversal of studying moving. It is a study of long-term staying and what creates long term staying. We show that family stability, in the sense of preserving family status as a couple or a single-person household, is strongly related to staying. It is also strongly related to the role of children in these households. The role of children is an important additive dimension, and they play an important role in the process of staying. We were able to show that households who had children at the beginning and the end of the window in the survey were very unlikely to move and even more unlikely to move if they were owners. For couple households in which the child entered after the first wave, there were also significant probabilities of staying. In contrast for single households, it was the presence of children at both the beginning and end of the survey that was important. The entry of a child stimulated the end of staying and a residential move. This is the classic adjustment related to family needs for space and neighborhood status.

As we hypothesized tenure is an important associated variable in long-term staying. But, note, unlike the move from rent-to-own and consequent mobility which is the way in which tenure is used in explaining mobility, here we focus on the association of long-term staying. It is strongly associated with staying as measured by the log odds. While tenure is strongly associated with staying, job change disrupts staying. Those with a job change are making residential moves to address relocation in response to employment transitions, and they are only half as likely to stay as those without a job change. Satisfaction did add moderately modestly to the explanation but is clearly subsumed in other measures. This chapter confirms the cross-sectional findings of tenure but provides important new findings on how children play a role in the duration of stay. It also provides a much more nuanced discussion of the nature and extent of duration across a longitudinal sample and how a subset of long-term staying can be explained by the life course evolution.

Acknowledgments This chapter uses unit record data from the Household, Income and Labour Dynamics in Australia (HILDA) survey. The HILDA project was initiated and is funded by the Australian Government Department of Social Services (DSS) and is managed by the Melbourne Institute of Applied Economic and Social Research (Melbourne Institute). The findings and views reported in this chapter, however, are those of the author and should not be attributed to either DSS or the Melbourne Institute.

Appendix

Table 8.7 Construction of variables used in the analysis

Category	xwaveid _hhpxid	Family type and stability	
Stable couple	Both members are each other's partner (_hhpxid) for all waves		
Stable single	Missing _hhpxid for all waves		
Transitioning single	Missing _hhpxid initially, then same for remaining waves		
Transitioning couple	Both members are each other's partner initially, then missing _hhpxid for remaining waves		
Other transitioning	All others		
Stayer	_mhli	Changed address during waves 1–13	For either member of a couple
Did not change address	0	In every wave	
Changed at least once	1	In at least one wave	
Demography			
Age	ahgage	Age as of last birthday, as of June 30, 2001	The older of the members of a couple
Single individual gender	ahgsex	Sex of single individual	Missing for a couple
Presence of children	ahhrih mhhrih	Child in household, in waves 1 and 13	
In wave 13	4, 12 1, 2, 3, 5, 6, 7	ahhrih mhhrih	
In wave 1	1, 2, 3, 5, 6, 7 4, 12	ahhrih mhhrih	
In waves 1 and 13	1, 2, 3, 4, 6, 7	Both ahhrih and mhhrih	
In neither waves 1 nor 13	4, 12	Both ahhrih and mhhrih	

(continued)

Table 8.7 (continued)

Highest education	_edhigh1	Highest education level achieved, in the earliest wave with a non-missing value	The highest of the members of a couple
BA or more	1, 2, 3		
Diploma	4		
HS graduate or certificate	5, 8		
Less than HS graduate	9, 10		
Status			
Occupation	_jbmo61 _esbrd	Occupation category, in the earliest wave for which the individual is employed	For a couple, professional if either is, otherwise technical/service if either is, etc.
Professional	_jbmo61: 1, 2		
Technical/service	_jbmo61: 3, 4		
Clerical/sales	_jbmo61: 5, 6		
Laborers/operators	_jbmo61: 7, 8		
Not employed	_esbrd: 2, 3	In every wave	
Household income	ahifefp ahifefn	Household financial year gross regular income ($) in wave 1	
Grouped into quintiles			
[see note 1]			
Household income change	ahifefp ahifefn mhifefp mhifefn	Change in household financial year gross regular income ($) between waves 1 and 13	
Grouped into quintiles			
[see note 2]			
Housing tenure	ahstenr	Household tenure in wave 1 [see note 3]	
Owner			
Renter			

Job change	_pjsemp _pjmsemp _pjemply	Changed employer (or primary employer) during waves 1–13	For either member of a couple
Changed at least once	[see note 4]	In at least one wave	
Did not change employers	Otherwise	In every wave	
Neighborhood advantage/ disadvantage	ahhad10	SEIFA 2001 index of relative socioeconomic advantage or disadvantage in wave 1	Deciles collapsed into quintiles
SEIFA quintiles			
Satisfaction			
Housing satisfaction	_losathl	Satisfaction with the home in which you live, averaged over all waves	For both members of a couple
Collapsed into quintiles			
[see note 5]			
Neighborhood satisfaction	_losatnl	Satisfaction with the neighborhood in which you live, averaged over all waves	For both members of a couple
Collapsed into quintiles			
[see note 5]			

Note 1: Quintiles of household income were computed across all households in wave 1 of HILDA. These were then merged back to the individual observations, so that the quintile definitions are independent of the selection of individuals for modeling universe

Note 2: For each household in the modeling universe, the difference between income in wave 13 and in wave 1 (in constant dollars) was calculated; then quintiles of this change were computed separately in each category (established couples and established singles)

Note 3: The coding of the tenure question changed in the second wave with the separation of rent-to-buy schemes from rental. We recoded the first wave response (ahstenur) to be consistent with the subsequent waves, imputing rent-to-buy for renters who were asked the value of their house (ahsvalue). In our models, rent-to-buy was combined with ownership, and living rent free or with life tenure was combined with rental or paying board

Note 4: The individual was considered to have changed employers if he or she was either (a) unemployed at the previous interview (_pjemply = 2) and employed at the current interview (_esbrd = 1) or (b) employed at the previous interview (_pjemply = 1) and not working for the same employer at the current interview (_pjsemp = 2 or _pjmsemp = 2)

Note 5: Quintiles of average satisfaction were computed separately for each satisfaction measure (household and neighborhood) in each category (stable couples and stable singles)

References

Burbank J, Keely L (2015) Baby boomers and their homes. Demand (Nielsen)

Clark WAV, Davies-Withers S (1999) Changing jobs and changing houses: mobility outcomes of employment transitions. J Reg Sci 39:653–673

Clark WAV, Dieleman F (1996) Households and housing: choice and outcomes in the housing market. State University, Center for Urban Policy Research, Rutgers

Clark WAV, Huff JO (1977) Some empirical tests of duration-of-stay effects in intraurban migration. Environ Plan A 9:1357–1374

Clark WAV, Lisowski W (2017) Prospect theory and the decision to move or stay. Proc Natl Acad 114:9505–9509

Clark WAV, Duque R, Palmares I (2015) Place attachment and the decision to stay in the neighborhood. Popul Space Place 23. https://doi.org/10.1002/psp.2001

Cohn D, Morin R (2008) American mobility, who stays put, where's home. Pew Research Center Demographic Trends, Washington, DC

Cooke T (2013) All tied up: tied staying and tied migration within the United States 1997–2007. Demogr Res 29:817–836

Coulter R (2013) Wishful thinking and the abandonment of moving desires over the life course. Environ Plan A 45:1944–1962

Coulter R, Van Ham M, Findlay A (2015) Rethinking residential mobility: linking lives through time and space. Prog Hum Geogr:1–23

Fernández-Carro C, Evandrou M (2014) Staying put: factors associated with ageing in One's 'lifetime home'. Insights from the European Context. Res Ageing Soc Policy 2(1):28–56. https://doi.org/10.4471/rasp.2014.02

Fischer PA, Malmberg G (2001) Settled people don't move: on life course and (im-) mobility in Sweden. Int J Popul Geogr 7:357–371

Goodman A (2002) Estimating equilibrium housing demand for 'stayers'. J Urban Econ 51:1–24

Goodman AC (2003) Following a panel of stayers: length of stay, tenure choice, and housing demand. J Hous Econ 12:106–133

Hanson S (2005) Perspectives on the geographic stability and mobility of people in cities. Proc Natl Acad Sci USA 102:15301–15306

Hjalm A (2014) The stayers: dynamics of lifelong sedentary behavior in an urban context. Popul Space Place 20:569–580

Ioannides Y (1987) Residential mobility and housing tenure choice. Reg Sci Urban Econ 17:265–287

Lewicka M (2011) Place attachment: how far have we come in the last 40 years. J Environ Psychol 31(3):207–230

Michelin F, Mulder CH, Zorlu A (2008) Family events and the residential mobility of couples. Environ Plan A 40:2770–2790

Morrison PA (1967) Duration of residence and prospective migration: the evaluation of a stochastic model. Demography 4:553–561

Morrison PA (1970) Movers and stayers: an analysis based on two longitudinal data files. Rand Paper

Morrison PS, Clark WAV (2016) Loss aversion and the duration of residence. Demogr Res 35:1079–1100

Mulder CH (1993) Migration dynamics a life course approach. Amsterdam, Thesis Publishers-PDOD Publications

Mulder CH, Malmberg G (2014) Local ties and family migration. Environ Plan A 46:2195–2211

Osterland A (2015) For retirees there is no place like their own home. CNBC May 19 2015

Summerfield M, Freidin S, Hahn M, Li N, Macalalad N, Mundy L, Watson N, Wilkins R, Wooden M (2014). HILDA user manual – release 13. Melbourne Institute of Applied Economic and Social Research, University of Melbourne

Taylor P, Morin R, Cohn D, Wang W (2008) Who moves, who stays put, where's home? Pew Research Report, Washington, DC

Uzzell D, Pol E, Badenas D (2002) Place identification, social cohesion, and enviornmental sustainability. Environ Behav 34(1):26–53
Watson N, Wooden M (2004) Assessing the quality of the HILDA survey wave 2 data. HILDA project technical paper series no. 5/04, Melbourne Institute of Applied Economic and Social Research, University of Melbourne
Woldoff RA (2002) The effects of local stressors on neighborhood attachment. Soc Forces 81(1):87–116

Chapter 9
Short-Term Relocation Versus Long-Term Migration: Measuring Income Transfers by Inter-provincial Employees Across Canadian Provinces

K. Bruce Newbold

Abstract Inter-provincial employees (IPEs), or individuals who work in one province and reside in another, have emerged as the main source for inter-provincial worker mobility within Canada, with their numbers far exceeding the number of inter-provincial migrants (IPMs, or individuals who permanently relocate from one province to another) on a yearly basis. Given the magnitude of the movement of IPEs, considerable amounts of income will be earned in one location and transported back to the place of residence. This is in contrast to IPMs, where income is earned and kept in the same place of residence/work. This chapter provides an estimate of the amount of income that is transferred across space by IPEs. Income-based versions of demographic effectiveness (Plane, Int J Popul Geogr 5:195–212, 1999) are applied to evaluate the movement of earned income in the Canadian context among inter-provincial employees, providing an estimate of the amount of income moved across space. Results illustrate the potential scale of income moved across space and the role of IPEs in redistributing income.

Keywords Canada · Migration · Income · Mobility

9.1 Introduction

Inter-provincial employees (IPEs), or individuals who work in one province and reside in another, have emerged as the main source for inter-provincial worker mobility within Canada, with their numbers far exceeding the number of inter-provincial migrants (IPMs, individuals who permanently relocate from one province to another) on a yearly basis. Driven in part by the growth of Canada's resource

K. B. Newbold (✉)
School of Geography & Earth Sciences, McMaster University, Hamilton, ON, Canada
e-mail: newbold@mcmaster.ca

R. S. Franklin (ed.), *Population, Place, and Spatial Interaction*, New Frontiers in Regional Science: Asian Perspectives 40,
https://doi.org/10.1007/978-981-13-9231-3_9

sector, the number of inter-provincial employees has grown by nearly 110,000 between 2004 and 2008, when the number of inter-provincial employees reached 453,000 yearly movers before declining slightly to 412,000 in 2009. In comparison, Statistics Canada counted 259,234 inter-provincial migrants in 2009. While a portion of inter-provincial employees includes the exchange of workers between Ontario and Quebec, and specifically between Ottawa and Gatineau, 37% of inter-provincial employees from the Atlantic provinces worked in Alberta in 2007, reflecting the growth of its energy sector and concomitant economic opportunities. As such, IPEs represent a significant number of workers and play an increasingly important role in the Canadian labor market, enabling individuals to respond to skill shortages and job opportunities over both the short and long term. Moreover, while IPEs and IPMs are broadly similar in terms of their sociodemographic profiles, research reveals subtle differences between these two groups, including differences in the age migration schedule and other sociodemographic characteristics of inter-provincial employees (Newbold 2018).

Beyond adjustments to labor supply and demand, the movement of inter-provincial employees also means that there is a substantial movement of income across provincial borders and more precisely between the municipality where the income was earned and the municipality where it is spent. In other words, IPEs will import or export income into or out of a region and its economy. In comparison, migrants earn income (work) in essentially the same location as they live. Not surprisingly, for example, smaller provinces draw a larger proportional share of wages from inter-provincial work. Morissette and Qiu (2015), for example, note that 8.5% of wages earned in 2011 by all employees residing in Newfoundland and Labrador came from inter-provincial employment, while the corresponding values for inter-provincial employees that call Quebec or Ontario home were much lower. However, the question that this chapter addresses is how much income is transferred across space by inter-provincial employees and what provinces and/or municipalities are the major senders or recipients of this income.

By applying Plane's (1999) technique to estimate income flows and their effectiveness across space, this chapter examines how inter-provincial employees redistribute income. Drawing on data from Statistics Canada's 2011 National Household Survey (NHS) and using methods first proposed by Plane (1999), income-based versions of demographic effectiveness are applied to evaluate the movement of earned income among inter-provincial employees, providing an estimate of the amount of income moved across space. Results reveal that substantial amounts of income are transferred across space. The balance of the chapter is structured as follows. Section 2 provides further insight into IPE movement and behaviors. Section 3 discusses the methods and data used in the analysis, and Section 4 presents the descriptive results. Section 5 concludes this chapter.

9.2 Background

Inter-provincial migration has provided an important, long-term labor adjustment mechanism that has redistributed the Canadian population and enabled regional and national economic growth (Finnie 1999). Over the recent past, however, the number and rate of individuals engaged in inter-provincial migration has declined, while the number of inter-provincial employees has increased, facilitated by technological advances that have made such short-term relocations more feasible, easier, and less expensive. Since the early 2000s, for example, the number of inter-provincial employees grew by over 110,000 to some 453,000 yearly movers in 2008 before dipping to 412,000 in 2009 following the recession. For those engaged in inter-provincial employment, the need to commit to the personal and familial costs of migration is removed, with workers instead cycling back and forth between their province of residence and province of work. While a portion of these movements are between neighboring labor markets such as Ottawa, Ontario, and Gatineau, Quebec (the constituent parts of Canada's capital region), a large number work in locations that are more distant from their residence. For example, 37% of inter-provincial employees from Atlantic Canada worked in Alberta in 2007, reflecting employment opportunities in the resource sector (Turcotte and Weeks 2014).

In general, the movement of inter-provincial employees across space is similar to that of inter-provincial migrants, with a greater number of exchanges between proximate regions. Still, there are also a significant number of individuals that travel long distances to work. Although inter-provincial employees and migrants are broadly similar in terms of their sociodemographic and socioeconomic profiles, they also differ in subtle but important ways (Newbold 2018). For instance, the migration age schedule for inter-provincial employees peaks slightly earlier (ages 20–24) than inter-provincial migrants (which peaks during the ages of 25–29) but remains relatively constant through the 30s, 40s, and 50s, whereas migration age schedules for migrants decline with increasing age (Rogers and Castro 1981). Subtle differences between IPMs and IPEs are also observed with respect to sociodemographic and socioeconomic effects, including differences in propensities by gender, educational attainment, industry, occupation, and home ownership. Overall, inter-provincial employees are more likely to be males and those engaged in occupations and industries including those in the resource sector, construction and transportation (Laporte et al. 2013; Morissette and Qiu 2015; Newbold 2018). At the same time, an increasing proportion of married households engage in inter-provincial employment, with women more likely to work in the accommodation or food services sector as well as wholesale and retail trade (Laporte et al. 2013).

Given the scale of inter-provincial employee movement, it is reasonable to assume that a large proportion of the earned income is exported from the place where it is earned and imported into the place of residence of the employee. Places with net losses of income could be seen as providing income subsidies to places with net in-migration, as observed in the United States (Nelson 2005). This movement of income across space can be evaluated through measures analogous to net migration and migration effectiveness (Plane 1999), with the methods used by Manson and Groop (2000) who documented the transfer of income down the urban hierarchy and resulting in greater income disparity in the United States and Newbold (2008) who explored the movement of non-earned income (i.e., pensions or other benefits) among older, nonworking migrants in Canada. Using data from Statistics Canada, the following explores the volume of income moved across space by inter-provincial employees.

9.3 Data and Methods

In order to estimate the amount of income moved between regions by inter-provincial employees, data is drawn from the "master" data file of Statistics Canada's 2011 National Household Survey (NHS[1]). The NHS allows the identification of IPEs given it records both their place of residence and the place of work (recorded at both the provincial and municipal level), information that is not typically found in other conventional labor mobility statistics which rely on a change in residential address (Laporte et al. 2013). The NHS also provides other sociodemographic and socioeconomic detail, including age, the presence of paid employment, income, and whether the respondent resided in an institutional home.

The research highlights three groups, including stayers, inter-provincial employees (IPE), and inter-provincial migrants (IPM) based on the province of residence in 2010 and 2011 and province of work in 2011. Stayers are defined as individuals who did not move and had the same province of residence and work in 2010 and 2011. Inter-provincial employees are individuals who are resident in one province in 2011 but who reported a place of work in another province in 2011. Inter-provincial migrants are defined as those whose province of residence on census day in 2011 differs from their province of residence 1 year prior (2010). All groups will be constrained to noninstitutionalized workers in the labor force aged 20–64 as of census day and who reported paid employment in 2010 (1 year prior to the NHS). Institutional residents are excluded from the analysis.

Two geographic scales will be considered, including the provinces of work and residence, as well as the top ten sending and receiving sub-provincial units (municipalities), defined by census metropolitan areas (CMAs), census areas (CAs), and rural (non-urban) remainders of each province. In the latter case, the top ten sending and receiving regions are defined in terms of the total amount (volume) of income gained or lost from the community. In addition, IPEs moving between proximate provinces but remaining within the same labor market (i.e., individuals from Gatineau, QC, working in Ottawa, ON, and vice-versa, as well as exchanges between Lloydminster Alberta and Lloydminster Saskatchewan) were removed from the

[1]The NHS data file was accessed through Statistics Canada's Research Data Centre (RDC) at the author's institution.

analysis, enabling the work to focus on inter-provincial employees that have moved a longer distance.

We are interested in the movement of income by IPEs and IPMs, with two measures used to assess this movement. First, net income migration is the difference between aggregate incomes of in-migrants and out-migrants or the increase (decrease) in aggregate provincial income associated with migration. Second, following Plane (1999), income effectiveness is defined as the ratio between net migration (in-migration – out-migration) and gross migration (in-migration + out-migration) income flows[2]:

$$E_Y = 100\left(\frac{Y_{in} - Y_{out}}{Y_{in} + Y_{out}}\right) \tag{9.1}$$

where Y_{in} represents the income of in-migrants and Y_{out} represents that of out-migrants. Negative values indicate that migration serves to remove income from a region, and positive values suggest that income enters a region. Values close to zero indicate no net change in income distribution. Demographic effectiveness is also presented in the following tables and is similarly defined to income effectiveness but with the number of in- and out-migrants to a province/region replacing measures of the income moved in and out of provinces and regions.

Clearly, there are caveats associated with the estimation of the movement of income by workers from one location to another. For the worker in Fort McMurray, Alberta, but resident in Goose Bay and Newfoundland and Labrador, for example, not all of their earned income will be carried back to Goose Bay. Instead, a portion of their income would remain in Fort McMurray in the form of rent, food, entertainment, and other expenses. Estimating the dollar value that remains in the location of work would not be a simple procedure without more detailed data, particularly without data associated with accommodations, as some workers will live in industry housing, while others will share housing. Similarly, transportation costs between residence and workplace may or may not be fully (or partially) paid by employers. Consequently, the current analysis assumes that the full amount of the reported earned income returns home, which ultimately overemphasizes the true value.

9.4 Results

In total, inter-provincial employees were responsible for the movement of nearly four billion dollars in 2010—a figure which excludes those moving between proximate labor markets. Table 9.1 lists the per capita income levels of IPEs, including income exported from the province of work (i.e., IPEs working in province X, where

[2]Additional information on these measures are found in Plane (1999), and the interested reader is referred to this.

Table 9.1 Total (per capita) income ($), inter-provincial employees aged 20–64, 2010

Province	Exported income from province of work ($)	Imported income to province of residence ($)	Stayers ($)	Exported/ stayers (%)	Imported/ stayers (%)
NL	55,865	51,540	45,260	123.4	113.9
PEI	46,475	35,520	41,225	112.7	86.2
NS	55,490	46,490	44,495	124.7	104.5
NB	52,170	42,405	43,000	121.3	98.6
QC	52,610	58,960	47,805	110.1	123.3
ON	55,645	58,085	57,820	96.2	100.5
MB	59,630	48,645	47,585	125.3	102.2
SK	59,225	57,980	50,960	116.3	113.8
AB	67,225	64,675	65,625	102.4	98.6
BC	68,645	55,295	52,785	130.0	104.8
Terr	58,800	74,780	64,020	91.8	116.8
Average	58,565	58,565	53,680	114.0	105.7

X is not their province of residence (column 1)), imported income (i.e., the income associated with IPEs and their province of residence (column 2)), and non-migrants (stayers) for each of the ten provinces and an aggregate territories unit that includes all three Canadian territories (Nunavut, Yukon, and Northwest Territories).

Generally, inter-provincial employees reported higher incomes than individuals who did not work across provincial boundaries (stayers) ($58,565 versus $53,680[3]).

Inter-provincial employees working in Alberta (AB) and British Columbia (BC) reported the highest per capita incomes ($67,225 and $68,645, respectively) that were then exported out of the province of work and into the province of residence. Conversely, inter-provincial migrants working in Prince Edward Island (PEI) and residing in another province reported the lowest exported per capita income ($46,475). Turning to the importation of income, IPEs returning to the Territories after working in another province reported the highest income ($74,780), with IPEs returning to PEI returning with the lowest reported income ($35,520). As such, it is likely that participation in the labor market as an IPE is not necessarily to maximize income returns. In the case of IPEs returning to PEI, for example, reported incomes are less than those of stayers, suggesting that the decision to work outside the province of residence is driven in part by the availability of jobs, along with the skills required in the position and the skills embodied in the worker both in the province and elsewhere. In other cases, participation in the labor market as an IPE results in significantly higher incomes than stayers.

The final two columns in Table 9.1 provide perspective on the relative income levels of inter-provincial employees and stayers. IPEs resident in PEI but working elsewhere have the lowest per capita incomes relative to non-migrants (86.2%), while IPEs resident in Quebec (QC) have the highest per capita incomes (123.3%). Further, IPEs returning to Newfoundland and Labrador (NL) (113.9%), Nova Scotia

[3]All dollar values are Canadian dollars as of census day in 2011.

Table 9.2 Net migration (N), net income migration ($), and income effectiveness: Inter-provincial employees aged 20–64, 2010

	Net migration (N)	Net income migration (thousands, $)	Demographic effectiveness (%)	Income effectiveness (%)
NL	−5540	316,079	−64.2	66.5
PEI	−770	47,003	−27.5	39.5
NS	−2105	149,907	−22.4	30.6
NB	−1950	133,814	−22.8	32.3
QC	−2825	110,578	−19.1	13.6
ON	−715	10,278	−2.9	0.7
MB	325	18,512	5.0	5.2
SK	−125	13,074	−1.3	2.4
AB	14,000	−885,182	47.0	−45.5
BC	−3675	344,326	−21.1	31.1
Terr	3375	−258,388	81.6	−85.3

(NS) (104.5%), Saskatchewan (SK) (113.8%), BC (104.8%), and the Territories (116.8%) have higher incomes relative to stayers, while the percentage is not different from 100 (i.e., equivalent incomes between non-migrants and IPEs) in the case of Ontario (ON). With reference to the province of work, inter-provincial employees working in all provinces but Ontario and the Territories had incomes typically much greater than non-migrants in their province of residence, reinforcing the income-generating opportunities available for those working outside their province of residence.

Table 9.2 reflects the movement of income and individuals across the country through measures of income and demographic effectiveness of income. The first column represents the net gain or loss of IPEs at the provincial scale. For instance, Alberta, Manitoba (MB), and the Territories hosted more IPEs than they sent to other provinces, while all remaining provinces sent more IPEs in search of employment than they gained. The second column captures the income equivalent of net migration or the difference between aggregate incomes of "in" and "out" inter-provincial employees (i.e., the total income that is transported between the earnings province and the province of residence, in thousands $). Returning to the example of Alberta, IPEs exported more income ($1.4 billion in 2010) from the province than returning IPEs entered with $530 million. Alberta's large net income loss reflects the large number of inter-provincial workers (21,890, with a net inflow of 14,000) it hosted in 2010. In addition to Alberta, the Territories had a net inflow of IPEs and were a net exporter of income to other provinces, while all other provinces experienced the net importation of income. As already noted, these values overrepresent the actual dollar value that was removed from (or imported to) a province, as the estimates do not account for expenses associated with such things as transportation, housing, entertainment, or food while working in away from their home province. Instead, the values should be treated as the upper maximum of what could be sent back to a province of residence.

The third column presents the demographic effectiveness measure, and the fourth column includes the income effectiveness measure, capturing the effects of net population gains (losses) and the differences in the per capita income of migrants (Plane 1999). The large net in movement of IPEs into Alberta was associated with a high demographic effectiveness (47.0%), while its net loss of income was reflected in negative income effectiveness (−45.5%). That is, such that while it experienced the net arrival of inter-provincial employees, the comparatively large volume of income flowing out of the province relative to that which was being imported by returning IPEs resulted in its large negative income effectiveness ratio. Newfoundland and Labrador, on the other hand, experienced a net outflow of inter-provincial employees. That is, with a demographic effectiveness of −64.2%, it sent more inter-provincial employees to other provinces than it received but had an income effectiveness of 66.5%, reflecting the relative scale of importing inter-provincial income as compared to its exportation of income.

Although only slightly greater than the net income imported into Newfoundland and Labrador $316 million), British Columbia had the largest net income inflow ($344 million) but more modest demographic and income efficiencies (−21.1% and 31.1%, respectively). Ontario, Canada's most populous province, had a modest positive net income of just $10 million in 2010 and an income effectiveness of just 0.7%, highlighting a relative equality in terms of both sending and receiving inter-provincial employees, as well as in terms of the income transferred in and out of the province, reflecting the province's role as both a sender and receiver of migrants.

Table 9.3 disaggregates IPE income flows using Plane's (1999) techniques, enabling greater understanding of provincial gains or losses associated with inter-provincial employee movement. This decomposition enables understanding of whether income gains or losses are driven by net inter-provincial employee movement (i.e., the volume of IPE movement into and out of a province) or by differences in the per capita income levels of IPEs. IPE streams are decomposed into the net migration component and the differential migration component, enabling the analysis to distinguish whether provinces gained or lost income through the volume or

Table 9.3 Typology of migration decomposition by province by inter-provincial employees aged 20–65: total income ($, thousands), 2010

	Net	Differential
NL	297,412	18,666
PEI	31,651	15,353
NS	107,534	42,372
NB	92,019	41,795
QC	157,483	−46,905
ON	40,488	−30,210
MB	−17,486	35,997
SK	7152	5922
AB	−923,152	37,971
BC	227,799	116,527
Terr	−225,356	−33,033

characteristics of migrants (Plane 1999). If the absolute value of the net migration component exceeds the value of the corresponding differential migration component, then the volume of migration is the major contributor to income redistribution. In cases where the differential migration component exceeds the net migration component, the characteristics of the migrants and the income they embody drive income redistribution.

Not surprisingly, the results presented in Table 9.3 highlight that net migration (as compared to the characteristics of migrants themselves) is driving the redistribution of income in almost all cases, reflecting results both in Plane (1999) and Newbold (2008). In other words, the volume of migration, and not the characteristics of IPEs, determined whether provinces had a net gain or loss of income. The one exception is Manitoba, where the differential component is greater, suggesting that there may be something particular about the composition of IPE flows in this instance.

Typically, the net migration component was larger than the differential component. In some cases, the two components worked in opposite directions, highlighted by Alberta with a net migration component of -\$923 million and a differential component of nearly \$38 million. In Ontario's case, the income differential (−\$30 million) explains why its demographic effectiveness was negative, while its income effectiveness was positive: despite a net loss of IPE workers (−712), it had a larger inflow of income as compared to outflow.

While provinces are a common scale to measure long distance migration, individuals are selecting particular destinations. Tables 9.4a and 9.4b therefore extend the analysis to the sub-provincial level, with sub-provincial areas including census metropolitan areas (CMA), census areas (CA), and rural (non CMA/CA) areas. Table 9.4a highlights the top ten income exporting regions (by total dollar value), and Table 9.4b highlights the top ten receiving regions (by total dollar value).

Turning first to Table 9.4a, Wood Buffalo, Alberta, a municipality in northern Alberta that is home to Fort McMurray and is the center for Canadian oil sands production, was an important destination for inter-provincial employees. In 2010, it had a net in-migration of 8405 workers, and it was the single largest exporter of income by IPEs (by volume), with approximately \$676 million exported by IPEs from the municipality in total during 2010. However, per capita income for IPEs working in Wood Buffalo was not the largest (\$58,550) and well below the per capita incomes of stayers (\$103,480). Other CMAs and CAs in Western Canada were also important places of work for IPEs (in terms of the total income exported), with Calgary, Edmonton (with both Calgary and Edmonton being cities with populations over 1 million), and rural Alberta also among the regions that exported large volumes of income in 2010, partially reflecting the resource-based provincial economies. But locations in Alberta were not the only large income exporters, with other areas including rural Ontario, rural British Columbia, and rural Saskatchewan also major exporters of income, potentially reflecting work in more remote locations or resource oriented firms with higher pay. In addition, the major CMAs of Toronto, Montreal, and Vancouver also exported large volumes of income.

Table 9.4a Per capita income transferred by IPEs, 2010: top ten exporting regions

	Total $ exported (thousands)	Exported per capita income from place of work ($)	Stayers ($)	Exported/ stayers (%)	Net migration	Demographic effectiveness (%)	Income effectiveness (%)
Wood Buffalo, AB	676,204	58,550	103,480	59.6	8405	93.6	−95.2
Calgary, AB	266,543	61,880	78,585	78.7	2025	32.8	−35.1
Toronto, ON	250,638	48,890	70,475	69.4	1155	21.0	−40.7
Edmonton, AB	203,193	66,235	63,230	104.8	1755	33.4	−27.2
Montreal, QC	185,294	62,435	53,140	117.5	575	9.9	−9.0
AB rural	181,888	77,860	48,725	159.8	1165	19.4	1.6
ON rural	170,398	60,070	43,385	138.5	−385	−5.5	13.4
BC rural	144,142	71,680	43,940	163.1	−260	−4.8	16.4
Vancouver, BC	143,429	65,665	60,315	107.2	−175	−3.9	1.6
SK rural	116,614	56,215	42,500	132.3	−530	−10.4	15.5

Table 9.4b Per capita income transferred by IPEs, 2010: top ten importing regions

	Total $ imported (thousands)	Imported per capita income to place of residence ($)	Stayers ($)	Imported/ stayers (%)	Net migration	Demographic effectiveness (%)	Income effectiveness (%)
NL rural	290,198	50,150	38,520	130.2	−4480	−70.0	71.6
ON rural	222,904	51,270	43,385	118.2	−385	−5.5	13.4
BC rural	200,757	56,685	43,940	129.0	−260	−4.8	16.4
QC rural	176,684	54,690	36,610	149.4	−2210	−39.9	32.2
AB rural	187,911	50,805	48,725	104.3	1165	19.4	1.6
Montreal, QC	162,976	61,300	53,140	115.4	575	9.9	−9.0
SK rural	159,269	50,655	42,500	119.2	−530	−10.4	15.5
Vancouver, BC	148,169	67,700	60,315	112.2	−175	−3.9	1.6
NB rural	143,986	43,555	36,565	119.1	−1275	−29.8	37.5
NS rural	122,039	49,440	37,840	130.7	−480	−11.9	16.8

Echoing results at the provincial level, demographic and income effectiveness typically work counter to each other, while flows of migrants are often unidirectional (with, e.g., Wood Buffalo receiving far more IPEs than it sends, generating a demographic effectiveness of 93.6%), the flow of income out of the region is equally unidirectional (income effectiveness = −95.2%). While Wood Buffalo was the most efficient, other regions also had strong unidirectional flows of income and IPEs.

Table 9.4b captures the major income receiving regions by total volume of income. While all of the regions have a net loss of IPEs (sending more than receiving), the large representation of rural locations, including rural Newfoundland, Quebec, Ontario, Saskatchewan, British Columbia, New Brunswick, Alberta, and Nova Scotia, among receiving regions is particularly striking. That is, a number of rural areas imported large volumes of income, implying that these same areas are a comparatively large source of IPEs. Moreover, IPEs are an important source of income for these rural areas, as workers find employment elsewhere. Rural Newfoundland and Labrador, for example, received over $290 million as IPEs returned home. Moreover, the per capita income of IPEs returning to these rural areas was consistently higher than stayers. For instance, IPEs that resided in rural Newfoundland and Labrador reported an income of approximately $50,150, while the reported income of their counterparts who did not engage in inter-provincial employment was just $38,520. As with sending regions, demographic and income effectiveness work counter to each other, with flows of workers corresponding to large flows of income in the opposite direction.

9.5 Conclusions

Inter-provincial employees represent a significant number of workers and are an important part of the Canadian labor market, with individuals responding to short- and long-term skill shortages and job opportunities across the country. Like inter-provincial migration, it would appear that inter-provincial employees help with macroeconomic adjustments to the Canadian labor market (Amirault et al. 2013; Carey 2014; Coulombe and Tremblay 2009; Courchene 1970, 1974; Newbold 2018; Rosenbluth 1987) as individuals temporally relocate to places with demand for employment and higher wages (Borjas et al. 1992). Importantly, however, inter-provincial employees are not forced to bear the social costs of relocating families or to give up location-specific capital such as housing and social networks in their place of residence, reflecting the endowment effect (Clark and Lisowski 2017).

With reference to both the provincial and sub-provincial scales, this chapter has explored how inter-provincial employees redistribute income across space, shifting it from the places where it was earned to places where they reside. Estimated to represent some $4 billion in 2010 alone, the movement of income by inter-provincial employees across the country is not insignificant. Not surprisingly, large volumes were exported from places that hosted a large number of IPEs, with these locations often representing rural or resource locales such as rural Alberta and Wood Buffalo, Alberta, home to Fort McMurray. More broadly, the province of Alberta was an important distributor of income. Locations with net losses of income could be seen as providing income subsidies to receiving regions. Conversely, the Atlantic provinces (inclusive of Newfoundland and Labrador, Prince Edward Island, Nova Scotia, and New Brunswick), as well as rural areas or locations with more limited economies and opportunities, were often the source locations for IPEs and benefited from the net inflow of income earned by these employees. Echoing results by Morissette and Qiu (2015), it would appear that smaller provinces and regions draw a larger proportion of wages from IPEs, with IPEs generally having higher per capita incomes than those who did not relocate for work, meaning that the movement of income is important for their economies. But, it is important to note that most areas (and particularly at the provincial level) both sent and received IPEs. While receiving a modest influx of income, the province of Ontario was a redistributor of income, neither benefiting nor losing from income transfers, reflecting the large number of workers that either work in the province and live elsewhere or live in Ontario and work elsewhere.

Two other findings should be noted. First, although participation in the labor market as an IPE often results in significantly higher incomes as compared to those who did not migrate, the results would appear to suggest that engaging in inter-provincial work is not just about maximizing income potential. In several cases, the per capita incomes of IPEs was less than that of stayers, suggesting that the decision to work outside the province of residence is driven in part by the availability of jobs, along with the skills required in the position and the skills embodied in the worker both in the province and elsewhere. Second, the movement of income between provinces and sub-provincial regions were predominantly associated with the volume of net employee movement, as opposed to differential income effects of IPEs as they move in and out of locations.

Although this chapter has provided estimates of the amount of income moved between places of work and residence, two caveats are noted. First, the 2011 NHS captured respondents prior to the 2014 downturn in the oil and gas sector which resulted in reduced construction of new facilities in Western Canada, decreased exploration, and increased automation in production and exploration. As a consequence, the downturn in the market and rationalization of employment needs has likely dampened the demand for inter-provincial employees. For example, Statistics Canada reported a loss of over 32,000 jobs between 2014 and 2017 in the mining, quarrying, and oil and gas extraction sector, with an additional loss of over 24,500 jobs in the related support sector in the same period. Furthermore, projections suggested that employment would not return to pre-2014 levels even as prices for oil rebound (Cattaneo 2016), while long-term economic prospects for investment and development in the oil sector remain dim. Consequently, the loss of jobs likely meant a reduced demand for IPEs over this period, with limited opportunities for growth in the post-2018 period. However, other opportunities associated with infrastructure projects may provide new opportunities for inter-provincial employee movement in the future. Second, as noted previously, the income values overrepresent the actual dollar value that was removed from (or imported to) a particular location, as the estimates do not account for expenses associated with such things as transportation, housing, entertainment, or food while working in away from their home province. Instead, the values should be treated as the upper maximum of what could be sent back to a province of residence.

References

Amirault D, de Munnik D, Miller S (2013) Explaining Canada's regional migration patterns. Bank Can Rev:16–28

Borjas GJ, Bronars SG, Trejo SJ (1992) Self-selection and internal migration in the United States. J Urban Econ 32:159–185. https://doi.org/10.1016/0094-1190(92)90003-4

Carey D (2014) Overcoming skills shortages in Canada. OECD economics department working papers, no. 1143, OECD Publishing. https://doi.org/10.1787/5jz123fgkxjl-en

Cattaneo C (2016) Oil and gas sector to shed 24,400 more jobs in 2016, says new report. The financial post, 21 April 2016. http://business.financialpost.com/commodities/energy/oil-and-gas-sector-to-shed-24400-more-jobs-in-2016-says-new-report. Accessed 25 Sept 2018

Clark WAV, Lisowski W (2017) Prospect theory and the decision to move or stay. Proc Natl Acad Sci USA 114(36):E7432–E7440. https://doi.org/10.1073/pnas.1708505114

Coulombe S, Tremblay JF (2009) Migration and skills disparities across the Canadian provinces. Reg Stud 43(1):5–18. https://doi.org/10.1080/00343400701654111

Courchene TJ (1970) Interprovincial migration and economic adjustment. Can J Econ 3:550–576

Courchene TJ (1974) Migration, income and employment. C. D. Howe Institute, Toronto

Finnie R (1999) The patterns of inter-provincial migration in Canada 1982–95: evidence from longitudinal tax-based data. Can Stud Popul 26(2):205–235

Laporte C, Lu Y, Schellenberg G (2013) Inter-provincial employees in Alberta. Statistics Canada, Catalogue no. 11F0019M — No. 350

Manson GA, Groop RE (2000) US intercounty migration in the 1990s: people and income move down the urban hierarchy. Prof Geogr 52:493–504. https://doi.org/10.1111/0033-0124.00241

Morissette R, Qiu H (2015) Interprovincial employment in Canada, 2002 to 2011. Ottawa: Statistics Canada, Catalogue no. 11-626-X — No. 047

Nelson PB (2005) Migration and the regional redistribution of nonearnings income in the United States: metropolitan and nonmetropolitan perspectives from 1975 to 2000. Environ Plann A 37:1613–1636. https://doi.org/10.1068/a37170

Newbold KB (2008) Inter-provincial migration and retirement income transfers among Canada's older population: 1996–2001. Environ Plann A 40:1501–1516. https://doi.org/10.1068/a39188

Newbold KB (2018) Short-term relocation versus long term migration: implications for economic growth and human capital change. Popul Space Place. https://doi.org/10.1002/psp.2211

Plane D (1999) Geographical pattern analysis of income migration in the United States. Int J Popul Geogr 5:195–212. https://doi.org/10.1002/(SICI)1099-1220(199905/06)5:3<195::AID-IJPG138>3.0.CO;2-K

Rogers A, Castro L (1981) Model migration schedules. Research Report-81-30, Laxenburg, Austria: International Institute for Applied Systems Analysis

Rosenbluth G (1987) The causes and consequences of interprovincial migration. Discussion Paper no. 87-29, University of British Columbia, Vancouver, Canada

Turcotte M, Weeks J (2014) The migration of infrastructure tradespersons. Statistics Canada, Catalogue no. 75-006-X

Chapter 10
Age Articulation of Australia's International Migration Flows

James Raymer, Nan Liu, and Xujing Bai

Abstract In this chapter, we are interested in how age profiles of migration vary by different immigrant groups arriving to or departing from Australia. Origin-destination patterns of international migration are examined, and a typology of age-specific migration is proposed, largely driven by the types of visas immigrants obtain to enter Australia. To form the typology, the study relies on observed immigration and emigration data from 1981 to 2016 for 19 different birthplace-specific groups, including the Australia-born population. The typology is utilised as a basis for examining how age profiles of international migration differ over time, by sex and across space. Our research shows how the types of visas used to enter the country may explain many of the differences in the observed age profiles of migration to and from Australia.

Keywords International migration · Age profiles · Migration policy · Australia

10.1 Introduction

Australia is a major immigrant receiving country with around 29% of its population born overseas in 2017 (ABS 2018). While the diversity of immigration streams to Australia by country of origin is well known (e.g. Hugo 2009; Jupp 2001; Khoo 2003; Markus et al. 2009; Richards 2008; Raymer et al. 2018; Wilson and Raymer 2017), such is not the case for the underlying age-specific patterns of migration. The ages of migrants entering (departing) represent an important aspect of demographic study (Castro and Rogers 1983; Plane 1993). They provide the basic information to understand demographic contributions to population age compositions and the numerical elements for studying levels of education, labour force participation,

J. Raymer (✉) · N. Liu · X. Bai
School of Demography, College of Arts and Social Sciences, Australian National University, Canberra, Australia
e-mail: james.raymer@anu.edu.au; nan.liu@anu.edu.au; xujing.bai@anu.edu.au

R. S. Franklin (ed.), *Population, Place, and Spatial Interaction*, New Frontiers in Regional Science: Asian Perspectives 40,
https://doi.org/10.1007/978-981-13-9231-3_10

fertility behaviours and retirement cohorts. They also capture the underlying life course motivations comprising the moves, which represent fundamental aspects of demography (Willekens 1999) and the study of migration (Bernard et al. 2014). In this chapter, we study the different age profiles of international migration coming to and departing from Australia by country or region of birth.

In thinking about how one might explain differences in the age patterns of migration, we draw from Plane and Heins' (2003) 'age articulation' article published in *The Annals of Regional Science* (see also Plane et al. 2005). Their paper examined inter-metropolitan migration flows in the United States during the 1985–1990 period with the aim of understanding the life course mechanisms driving the patterns. Through factor analysis, they identified seven clusters of age-specific migration that included college bound, leaving college, average, retirement, families and labour market, young families and labour market and older elderly. As stated in their article, age articulation refers to different shapes of age-specific migration found in origin-destination flows of migration.

Similar to Plane and Heins (2003), we rely on aggregate origin-destination migration flows by age groups but with international migration to and from Australia as our main interest. We are also interested in the effects of time, sex and locations on age articulation. In this chapter, we identify a typology of age-specific immigration and emigration to describe the main motivations and characteristics of migrants entering and departing Australia. This research is part of a larger Australian Research Council Discovery Project on 'the demographic consequences of migration from, to and within Australia'. The aim of the project is to understand the sources of population growth of different immigration streams on their long-term contributions to population change.

Age profiles of immigration and emigration are required for understanding the demographic effects of Australia's rapidly growing immigrant populations. Age profiles capture the contributions to the population's changing age composition, including the female immigrants at risk of producing children within Australia. Age profiles of migration also reflect the different life course motivations underlying the international moves. For example, flows of international students exhibit immigration and emigration age profiles with relatively high concentrations of people in their young adulthood years. Labour flows, by contrast, display age profiles with young adults joined, in many instances, by their spouses and children. Age-specific emigration patterns of overseas-born persons also exhibit peaks in young adult years but are often slightly older than the corresponding immigration streams and may include substantial numbers of persons around their retirement age (e.g. ages 60–69 years).

Apart from revealing life course motivations, age profiles of migration are useful for assessing the economic and social impacts of international migration on Australia, which in turn have profound implications for policymaking and projections. For instance, large inflows of working-age migrants expand Australia's labour force and help boost its economy, while sizeable inflows of children and the elderly may increase demand for social services, such as nurseries, primary schools and health care. Indeed, these issues are discussed in Plane and Heins' (2003, p. 109) paper, along with a basic framework that we apply in this study:

1. Use geographic units that capture the '. . .greatest meaning with respect to current economic, cultural, and social conditions that mediate movement behaviour'.
2. '. . . only when both the origin and destination of migrant streams are examined is it possible to establish properly the interregional differentials between economic, cultural, and social conditions that truly do affect the size of migration streams'.
3. '. . .stage-in-the-lifecycle is the first and foremost micro-scale predictor of migratory behaviour. . .'.

Considering the above framework, we focus on age-specific international migration flows defined by the top sending countries of birth and remaining "rest of region" birthplaces. Furthermore, we are interested in the demographic impacts of international migration on Australia's society over time, by sex and across states or territories within the country. In Australia, immigrants are highly concentrated in state capital cities and particularly in Sydney, Melbourne and Brisbane. Many areas outside these capital cities are in need of labour. Thus, government policies are in place to encourage migration to regional areas outside capital cities, albeit with limited success (Hugo 2008; Hugo and Harris 2011; Raymer and Baffour 2018).

10.2 Data

We focus on the age schedules or age profiles of migration, calculated by dividing each age- and birthplace-specific migration flow by the corresponding flow summed across age groups. This is in line with early research on analysing and parameterising age-specific migration by Rogers et al. (1978) and Rogers and Castro (1981).

The analyses in this chapter draw on two types of data: (1) reported and estimated flows of immigration and emigration flow data and (2) statistics on visa entries and exits. Data on immigration and emigration flows are sourced from the Australian Bureau of Statistics (ABS). They represent a time series of annual immigration and emigration from 1981 to 2016 disaggregated by age (0–4 years, 5–9 years, . . ., 80–84 years, 85+ years), sex, geography (8 states and territories of Australia) and place of birth (19 countries or regions). Since annual immigration and emigration flows by birthplace are only available from 2004 to 2016, annual birthplace-specific data on overseas arrivals and departures, based on visa category entries, were used to fill in the time series from 1981 to 2003. The methodology for producing the consistent time series of international migration data is described in Raymer et al. (forthcoming).

Flows of immigration and emigration to and from Australia, respectively, were small in the 1980s and 1990s for many of the overseas-born populations. For example, only around 7000 China-born immigrated to Australia during the 1981–1986 period with a corresponding 1300 persons emigrating. Similarly, around 6700 India-born persons immigrated to Australia in the early 1980s with 1500 persons emigrating. The issue of sparse data is even more acute at the state level. For instance, although more than 111,000 persons born in North Africa and the Middle East migrated to Australia during 2011–2016, only around 400 persons went to the Northern Territory. Out of the approximately 40,000 North Africa- and Middle East-born emigrants in the same period, less than 100 departed from Northern Territory. As another example, there were 138,000 persons born in Southern and Central Asia who migrated to Australia during the 2011–2016 period with just around 2100 settling in Tasmania.

The small flows of immigration and emigration at the national level in the early periods and the small states and territories in all periods of the study often resulted in irregularly shaped age profiles. Irregular age patterns of migration due to sparse data are not only hard to interpret but are also less useful when it comes to comparing age profiles over time, between birthplaces, and across space. A similar issue occurred in Plane and Heins' (2003) analysis of inter-metropolitan flows in the United States, where flows less than 100 persons were removed because of their "unreliable age profile" (p. 113). While there are methods for smoothing irregularly shaped age profiles of migration (see., e.g. Bernard and Bell 2015; Rogers et al. 2010), we focus on the more recent national level flows and visa statistics (described below) to develop the typology of age-specific migration.

The second source of data used in this chapter represents annual visa statistics by citizenship sourced from the Department of Home Affairs. It consists of (i) student visas from 2006 to 2016 by applicant type, sector and sex, (ii) temporary graduate visas by sex (2006 not available), (iii) working holiday visas by sex, (iv) temporary skilled visas by age and sex and (v) permanent visas by stream and sex. Note, detailed visa statistics by citizenship are not available for any years prior to 2006.

10.3 Typology of Age-Specific Migration

Age profile typologies are useful for simplifying and understanding the key differences present in migration data (see, e.g. Pittenger 1974, 1978; Plane and Heins 2003; Raymer and Rogers 2008). In this section, we focus on age profiles of immigration and emigration by country or region of birth during the most recent 5-year period, i.e. 2011–2016. The observed 19 age profiles exhibit several common patterns, indicating that migrants from certain birthplaces may come to Australia for similar purposes. To investigate the motivations underlying these 19 observed age profiles of immigration or emigration, we compare the patterns with visa statistics provided by the Department of Home Affairs in the financial years of 2011–2016. In doing so, we assume correlation between migration flows by citizenship and migration flows by country or region of birth.

The ABS immigration and emigration data are based on accounts of individual travel patterns that result in persons staying (departing) in Australia for 12 out of 16 months. To align our analyses of visa statistics with the ABS methodology for measuring international migration, we decided to exclude visitor visas from our analyses because of their short-term nature, i.e. usually less than 3 months (DIBP 2016c). Student visas granted for English Language Intensive Courses for Overseas Students (ELICOS) and other non-award sectors are also excluded because their durations of stay are often less than 1 year (DIBP 2016b). However, a portion of working holiday visa holders are included in our analyses because approximately 20% remain in the country for 12 out of 16 months (DIBP 2013, 2015, 2016a). Finally, humanitarian visas are not incorporated because of their relatively small numbers. In summary, the entry visa types used in our analyses to examine the motivations behind birthplace-specific immigration and emigration include (i) student visas (excluding the independent ELICOS sector and non-award sector), (ii) temporary graduate visas, (iii) working holiday visas (20%), (iv) temporary skilled visas and (v) permanent visas.

The five types of entry visas described above were further broken down into subcategories and then reclassified into four broader visa groups representing students, young labour, labour and family. The students group includes the primary applicants of student visas. The young labour group is comprised of temporary graduate visa holders, working holiday visa holders and temporary skilled visa holders aged between 15 and 29 years (DIBP 2016d). The labour group includes temporary skilled visa holders aged 30 years and above and permanent skilled visa holders. Finally, the family group consists of secondary applicants of student visas (restricted from studying more than 3 months and from working more than 40 h per fortnight), temporary skilled visa holders aged under 15 years, and permanent family visa holders (DIBP 2016b).

The relative shares of the regrouped entry visa types by country of citizenship in the periods of 2006–2011 and 2011–2016 are presented in Table 10.1. These shares are calculated from the detailed visa statistics provided in Appendix 1 and Appendix 2. Such statistics are not applicable for Australian citizens because they do not need visas to enter. Likewise, New Zealand citizens also have a special category visa that does not specify travel purposes, such as for study, work or permanent settlement (DIBP 2016a).

In examining the age patterns of migration and their corresponding visa breakdowns set out in Table 10.1 (see also Appendices 1 and 2), we settled on four main classifications of flows to form our typology: (i) primarily labour, (ii) mixed students and young labour, (iii) mixed labour and family and (iv) primarily students. In Fig. 10.1, the 19 birthplace-specific immigration and emigration flows observed during the 2011–2016 period are grouped into these four classifications.

In Fig. 10.2, six age profiles of immigration and emigration flows are presented. They represent the average age profiles for the four overseas-born migration flow classifications presented in Fig. 10.1 (summed across birthplaces) plus age profiles for the Australian-born and overall total populations. The *primarily students* classification exhibits the earliest young-adult peaks amongst the four classifications. Here, the peaks occur in ages 15–24 years for immigration and 20–24 years for emigration. These peaks suggest that young migrants come to Australia to commence their post-high school education and leave after they complete their degree in Australia. The *mixed students and young labour* classification, by comparison, is associated with higher and slightly later young-adult peaks. Reflected in these patterns are highly mobile young adults who have likely obtained their degrees somewhere else. They also appear to be short-term nature, as implied by the moderate differences between corresponding age peaks of immigration and emigration.

Table 10.1 Entry visa types (%) by country of citizenship, 2006–2011 and 2011–2016

	Student (primary applicant)		Temporary graduate working holiday temporary skilled (aged 15–29)		Temporary skilled (aged 30 and above) permanent skilled		Student (secondary applicant) temporary skilled (aged under 15) permanent family	
Citizenship	2006–2011	2011–2016	2006–2011	2011–2016	2006–2011	2011–2016	2006–2011	2011–2016
Australia	–	–	–	–	–	–	–	–
New Zealand	–	–	–	–	–	–	–	–
Other Oceania	30	38	3	5	24	20	43	37
United Kingdom	4	4	21	29	56	49	20	18
NW Europe	21	15	43	45	22	28	14	13
SE Europe	32	29	13	21	24	25	30	25
N. Africa and M. East	39	32	3	5	17	24	41	40
Vietnam	58	52	2	7	8	11	33	30
Philippines	12	16	5	9	51	46	32	29
Malaysia	56	52	4	10	28	25	12	13
Indonesia	57	55	4	9	15	12	24	23
SE Asia	58	53	3	6	14	14	25	27
China	61	57	4	10	22	19	13	14
NE Asia	45	38	26	34	15	14	14	14
India	43	23	10	19	28	40	19	19
SC Asia	48	33	5	14	23	30	24	23
North America	26	18	20	24	28	34	26	24
South America	53	47	5	10	17	20	25	23
Sub-Saharan Africa	21	25	5	9	49	40	25	27

Note: Calculations based on visa statistics obtained from the datasets provided by the Department of Home Affairs: https://data.gov.au/organization/immi

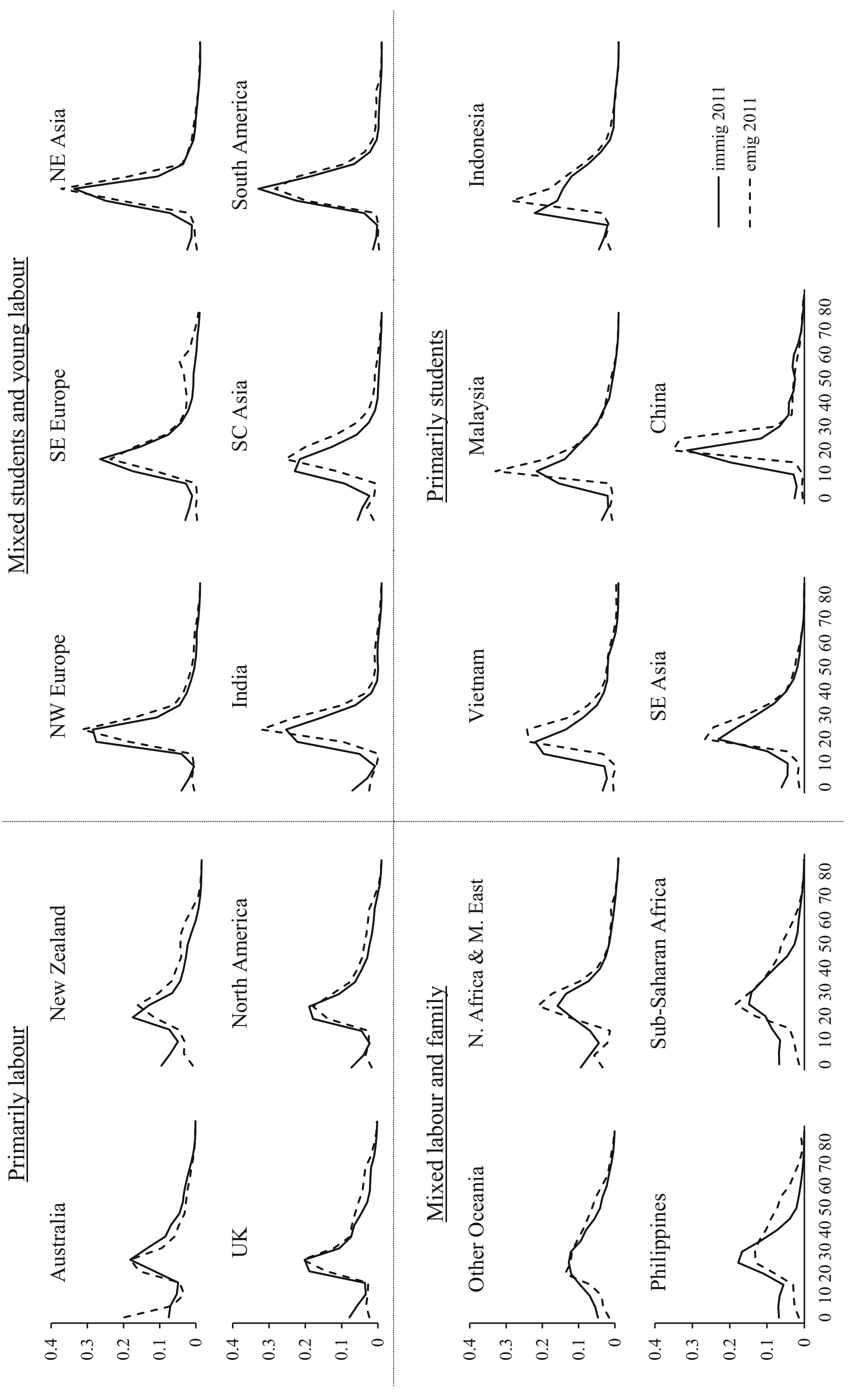


Fig. 10.1 Typology of immigration and emigration flows by country or region of birth, 2011–2016

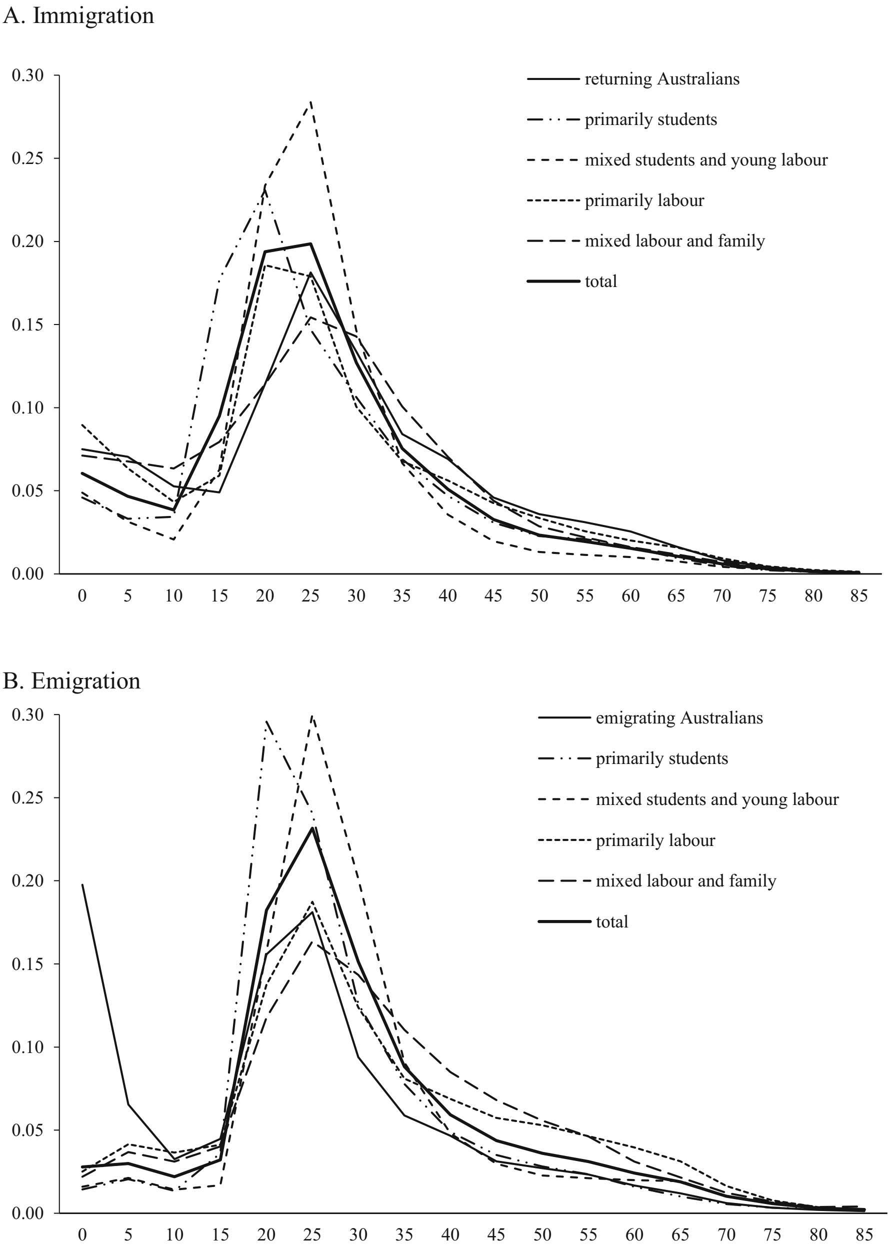


Fig. 10.2 Typology of age-specific immigration and emigration for Australia, 2011–2016

While both the *primarily students* and the *mixed students and young labour* classifications exhibit below-average proportions in ages 0–14 years, the *primarily labour* classification is characterised by above-average proportions in these age groups. The assumption is that these migrants bring their children with them.

Another notable feature is that, amongst the four classifications, it has the largest proportions of elderly or retirement-aged persons.

The *mixed labour and family* classification has the highest proportions of child-aged immigration. Another type of migration that is likely to be included in these flows are humanitarian migrants (i.e. asylum seekers and refugees). This is based on auxiliary information obtained from DIBP (2016a, b, c, d), which showed that seven out of the top ten countries receiving humanitarian support during the period of 2011–2015 belonged to either North Africa and the Middle East or Sub-Saharan Africa. In terms of emigration, the *mixed labour and family* classification contains higher-than-average proportions of outgoing persons aged 35–54 years.

10.4 Changes Over Time, by Sex and Across Space

In this section, we explore the consistency in the typology of age-specific international migration presented in the previous section over time, by sex and across states and territories in Australia.

10.4.1 Time

We first look at the changes in age profiles since 1981. Here, we selected four birthplaces for illustration, including Australia, New Zealand, the United Kingdom and China, and focused only on the immigration flows. As shown in Fig. 10.3, the evolution of the age profiles of the four birthplace-specific migration flows exhibits differing patterns. Steadily increasing from the early 1980s, the proportions of returning Australians in their young adulthood years reached their peaks in the early 2000s but dropped back to their starting level in the early 2010s. These changes were largely offset by the changes in the proportions of Australian children coming back to their birthplace, which experienced continuous decline until the period of 2001–2006.

New Zealand-born migration flows, meanwhile, exhibited an older age profile over time (Fig. 10.3). The proportions of New Zealand-born immigrants aged 20–24 years fell from 25% in the period of 1981–1986 to 15% in the period of 2001–2006 and stayed below 20% afterwards, while the proportions of incoming New Zealanders aged between 40 and 64 years experienced noticeable increases during these periods.

The age profiles of United Kingdom-born immigration flows, by contrast, became younger (Fig. 10.3). The increases in the proportions of United Kingdom-born migrants aged 20–29 years observed between the early 1980s and 1990s were accompanied by the decreases in the proportions of incoming United Kingdom-born children and persons in their 60s and their 70s, whereas such increases observed

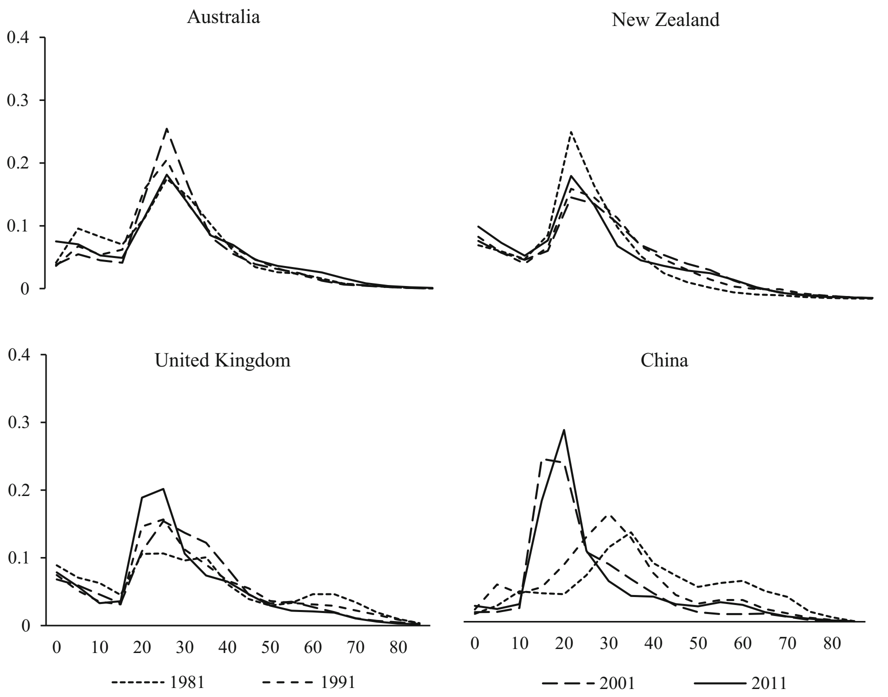


Fig. 10.3 Age proportions of immigration from selected birthplaces, 1981–1986, 1991–1996, 2001–2006 and 2011–2016

between the early 2000s and 2010s were accompanied by the drops in the relative numbers of United Kingdom-born immigrants aged between 30 and 44 years.

The age profiles of China-born immigrants also became younger over time (Fig. 10.3). This actually happened in a more dramatic way when compared to the shifts in the age profiles of United Kingdom-born immigrants. An important reason for such dramatic changes is the sparse data on China-born migration flows in the early 1980s and 1990s, which resulted in irregularly shaped age profiles of China-born immigrants in these two periods. Another noteworthy feature of the age profiles of China-born migrants is the low proportions of China-born children migrating to Australia over time, in stark contrast to the age profiles of the other three migrant populations.

We now examine changes in age profiles that have occurred over the past decade, with a particular interest in any transition in the typology of age-specific migration flows. In Fig. 10.4, the age profiles of immigration and emigration for the 19 birthplace-specific populations are presented for the 2006–2011 and 2011–2016 periods. In general, we find that the age profiles remained remarkably consistent over time with the exceptions of Northeast Asia-born and India-born migration. These two flows experienced transitions from *primarily students* in 2006–2011 to *mixed students and young labour* in 2011–2016. More specifically, the share of primary

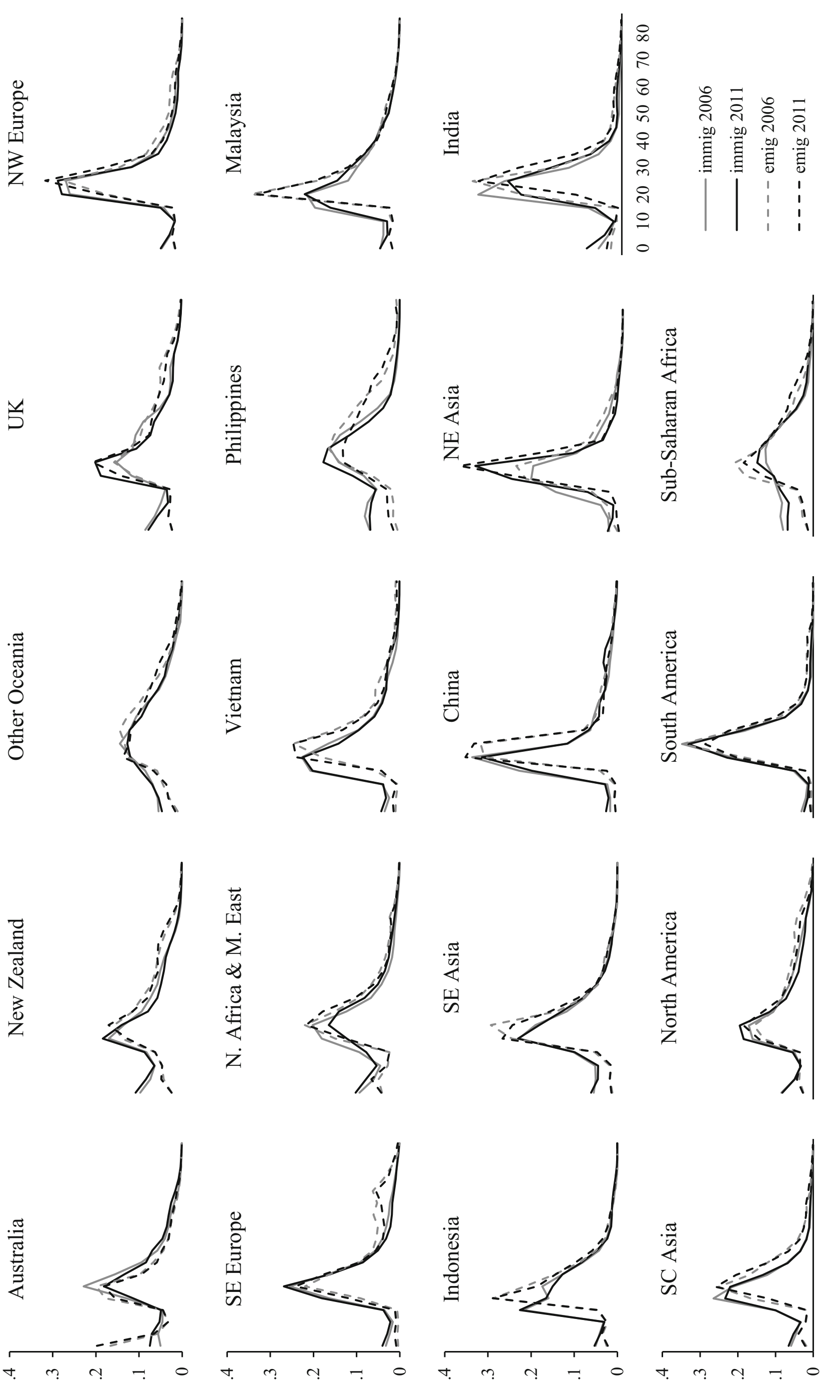


Fig. 10.4 Age proportions of immigration and emigration flows from 19 birthplaces, 2006–2011 and 2011–2016

student visa holders for the Northeast Asia-born immigration decreased from 45% in 2006–2011 to 38% in 2011–2016, while the share of visas issued to young labourers increased from 26% to 34%, respectively. The share of primary student visa holders for the India-born immigrant group witnessed an even sharper decline from 43% in 2006–2011 to 23% in 2011–2016, while the share of labour visas increased from 38% to 59%, respectively. The shift in the age profile of Northeast Asia-born immigration was driven by increases in the number of young adult labour visas, notably working holiday makers, whereas the change observed by India-born immigration was contributed by increased young adult labourers and decreased numbers of students. The significant decline in the number of India students was also likely due to a change in the immigration policy in 2010, which removed occupations, such as hairdressers and cooks, from the skilled migration occupation list (Parliament of Australia 2010).

Changes in the age profiles of migration of persons born in the United Kingdom and North Africa and the Middle East are also worth pointing out. Immigration of United Kingdom-born persons became younger between the two recent migration periods. This is a consequence of an increased young adult labour from 21% in 2006–2011 to 29% in 2011–2016. The factors contributing to this change include increased working holiday makers and young temporary skilled workers and decreased permanent skilled workers. These factors also help to explain the rises in the proportions of United Kingdom-born emigration of persons aged 25–29 years. By contrast, immigration of persons born in North Africa and the Middle East experienced reductions in the shares of young adults and growth in the shares of children and middle-aged persons. Detailed visa statistics reveal that while the number of student visas issued to applicants from North Africa and the Middle East remained more or less the same over the two most recent periods, there were more temporary skilled visas and permanent skilled visas granted in recent years.

Finally, another phenomenon worth mentioning pertains to the growing immigration of older persons from Vietnam and China. Looking closely at the age profiles of Vietnam-born and China-born immigration in Fig. 10.4, small yet noticeable upward shifts in the proportions of immigrants aged from 50–54 years to 65–64 years can be observed between 2006–2011 and 2011–2016. These immigrants are likely joining their family members already present in Australia. Curiously, this pattern is not observed amongst other immigrant groups.

10.4.2 Sex

In Fig. 10.5, we present the age profiles of immigration and emigration during the 2011–2016 period for the 19 birthplace-specific populations by sex. The most notable differences are found by persons born in the Philippines and North America. To further illustrate the differences for these two groups, consider the selected entry visa types for Filipino and North American citizens presented in Table 10.2 (see Appendix 3 for all the entry visa types). The share of North American males holding temporary skilled visas aged 30 years and above (28%) was 12 percentage points

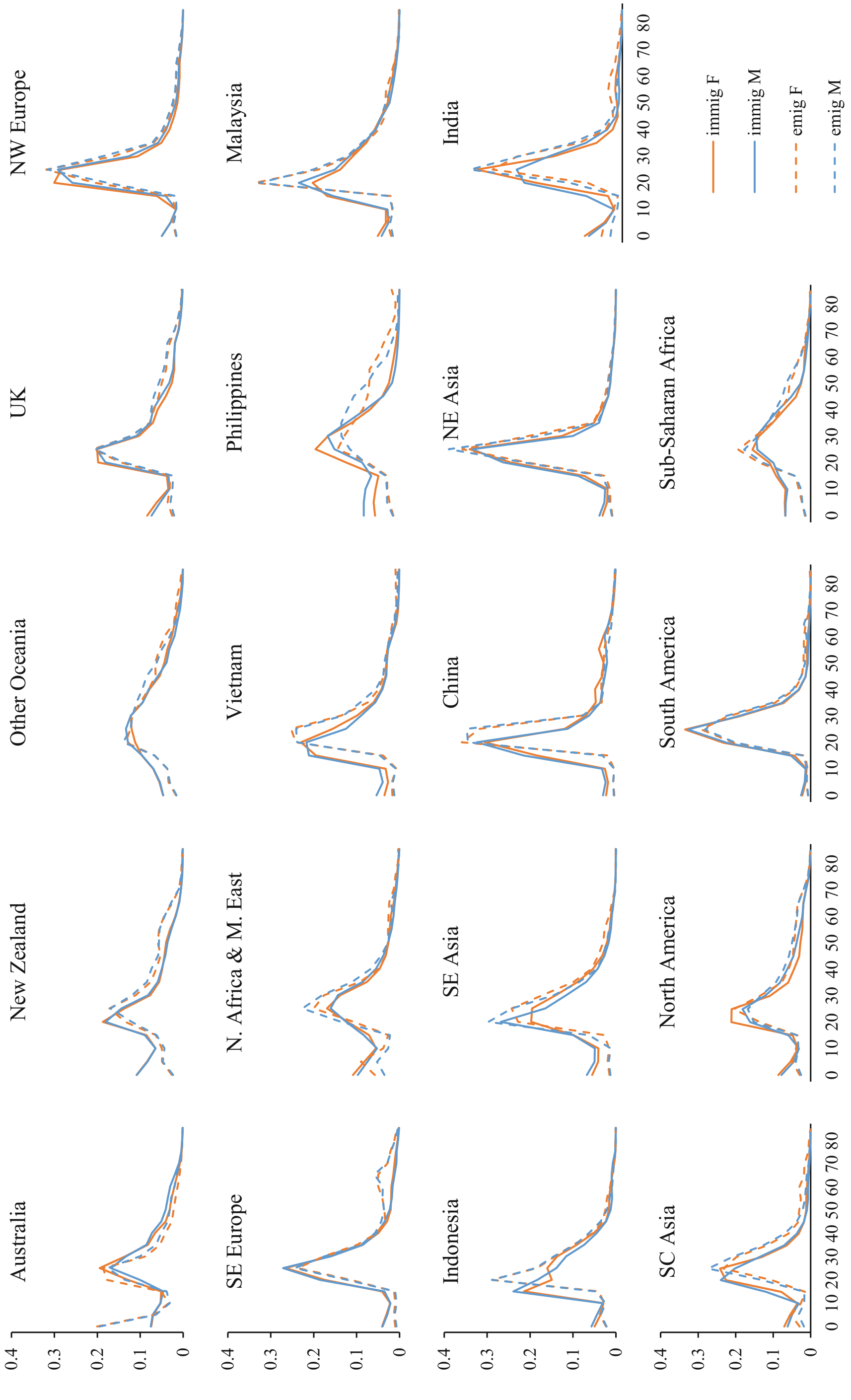


Fig. 10.5 Age proportions of immigration and emigration flows from 19 birthplaces by sex, 2011–2016

Table 10.2 Selected entry visa types (%) by sex for Philippines and North America citizenships, 2011–2016

Citizenship	Student (primary applicant)	Working holiday	Temporary skilled (aged 30 and above)	Permanent family
Philippines				
Male	11.90	–	19.82	8.55
Female	19.69	–	9.29	21.93
North America				
Male	15.66	11.74	40.46	10.84
Female	19.62	15.74	28.02	15.30

higher than that of North American females (16%), whereas the shares of North American females holding student visas and working holiday visas exceeded those of their male counterparts by 4 percentage points each. These patterns indicate that while the age profile of North American males belongs to the *primarily labour* classification, the age profile of North American females sits between the *mixed students and young labour* classification and the *primarily labour* classification.

Another interesting case of sex-specific age profiles of migration pertains to the Philippines-born population. Although the age profiles of both genders belong to the *mixed labour and family* classification, female immigrants from the Philippines are noticeably more concentrated in younger adult age groups than their male counterparts. Referring to the selected entry visa types in Table 10.2, the younger Filipino female migrants were mainly contributed by larger shares of student visa holders (20%) and permanent family visa holders (22%), in contrast to their male counterparts, who exhibited larger shares of temporary skilled visa holders aged 30 years and above. Auxiliary information on visa statistics has further shown that the number of spouse visas granted to female Filipinos was more than triple the number issued to male Filipinos for the year of 2012–2013 and around quadruple the number during 2014–2015 (Home Affairs 2018a, b). The family-oriented immigration patterns of Filipino females may be associated with the relatively high numbers of persons emigrating from Australia in their late 50s and their 60s for the purposes of returning home to care for ageing parents or to rejoin family.

10.4.3 Geography

Similar to comparing the typologies of migration by sex, the comparison of age profiles of birthplace-specific migrant groups across the eight states and/or territories in Australia focuses on the most recent 5-year period, i.e. 2011–2016. To simplify the comparison, we calculate age-specific ratios of the state- or territory-level age profiles to the national-level age profiles for each birthplace group and direction of flow. Ratios equal or near to one indicate similarity between the state or territory age profiles of migration flows and the national age profiles. Age-specific ratios larger (smaller) than one means that persons in those age groups are particularly attracted (not attracted) to that state or territory.

The results of the age-proportion ratios for all the 19 birthplace-specific migrant groups are shown in Appendix 4. We find age profiles of migration vary much more across geographic units than they do over time or by sex. Overall, age patterns of birthplace-specific migrant groups in New South Wales (NSW), Victoria (VIC) and Queensland (QLD) are similar to their respective age profiles at the national level. By comparison, age patterns in South Australia (SA), Western Australia (WA) and Australian Capital Territory (ACT) usually belong to the *primarily labour* classification or the *mixed labour and family* classification. Age profiles in both Tasmania (TAS) and Northern Territory (NT), by contrast, often resemble the *mixed students and young labour* classification, with most young labourers being working holiday makers, although age profiles in Tasmania also exhibit high proportions of elderly or retirement-aged migrants.

To illustrate changes in age-profile typologies across states and territories in more detail, four states (New South Wales, Victoria, South Australia and Tasmania) and four birthplaces (Australia, United Kingdom, China and India) are selected and presented in Fig. 10.6. Due to the strong positive correlations between immigration and emigration age-proportion ratios, we only present the immigration age-specific ratios. Here, we find that Australia-born immigration flows exhibit similar age patterns in New South Wales and Victoria, both mirroring their age profile at the national level. The age patterns of Australia-born immigration to South Australia and Tasmania, by comparison, do not resemble the national-level age profiles with considerably more middle-aged and retirement-aged persons.

Compared to Australia-born migrants, United Kingdom-born migrants exhibit more noticeable variations in their age profiles across states. The proportions of United Kingdom-born migrants aged 20–24 years and 25–29 years in New South Wales and Victoria are both higher than the proportions at the national level, indicating these states' relative attractiveness to these ages. South Australia, by comparison, features higher proportions of United Kingdom-born migrants both aged below 15 years and aged above 35 years. Here, the patterns resemble more the *mixed labour and family* classification than the *primarily labour* classification. Immigration to Tasmania of United Kingdom-born persons, on the other hand, attracts considerably more middle-aged labourers and retirees.

Varying age profiles across space are also observed for immigration of persons born in China and India. While the age profiles of China-born immigration to the four selected states all fit within the *primarily students* classification, this pattern is much more distinct in Tasmania than it is in New South Wales, Victoria and South Australia. Specifically, the proportions of Chinese immigrants aged 20–24 and 25–29 years in Tasmania were more than 1.5 times as high as the corresponding proportions observed at the national level. Another interesting pattern is that New South Wales exhibits higher proportions of immigration of Chinese-born persons aged 50 years and older. Immigration of India-born persons displays a particularly unique age profile in South Australia. This flow exhibits much higher proportions aged below 10 years and in the ages 30–44 years. These patterns indicate that the age

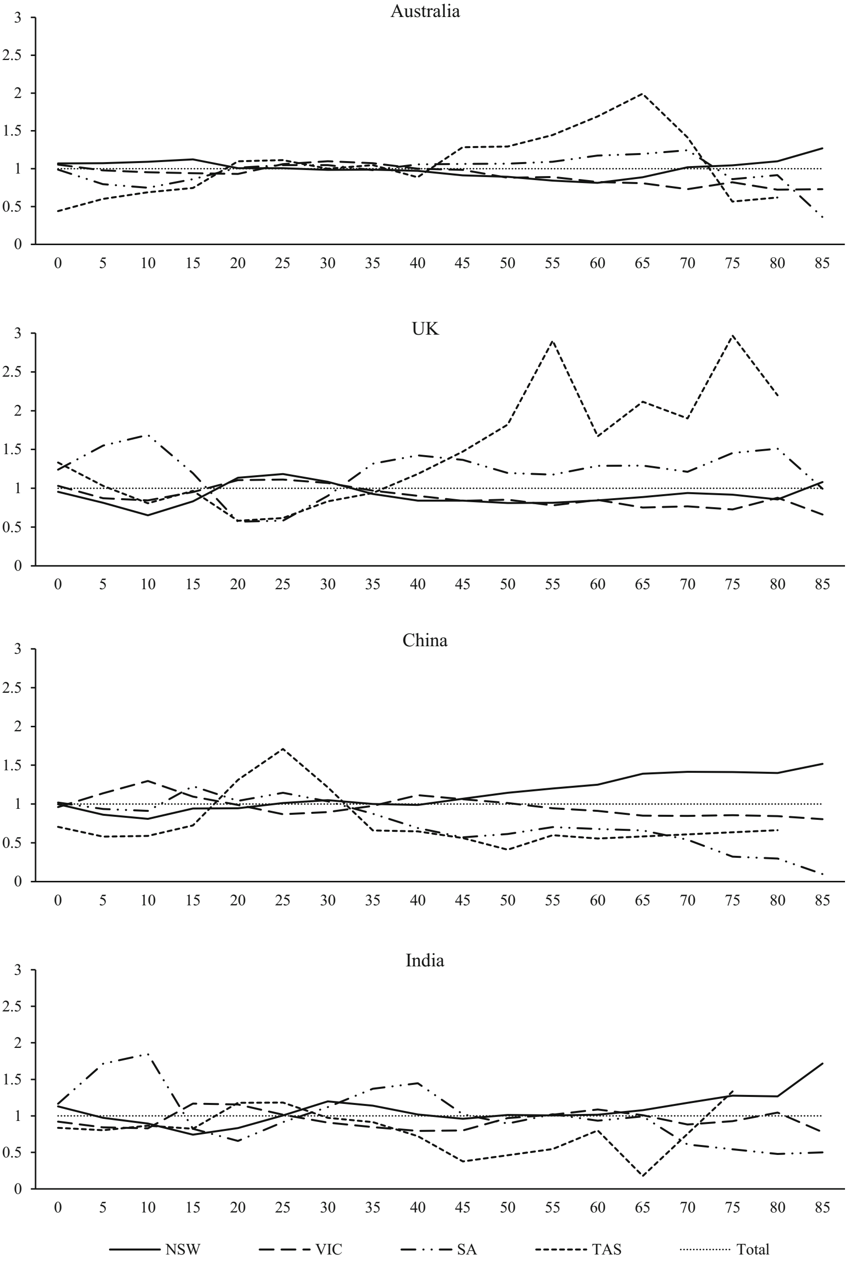


Fig. 10.6 Selected age-specific ratios of state or territory immigration age profiles to national age profiles, 2011–2016

profile of India-born migrants to South Australia resembles the *mixed labour and family* classification rather than the *mixed students and young labour* classification.

10.4.4 Summary

Age profiles of international migration to and from Australia for 19 birthplace-specific populations have been classified as *primarily students*, *mixed students and young labour*, *primarily labour*, and *mixed labour and family*. Using this typology as a basis, we have explored changes in the typology of age-specific migration flows over time, by sex and geography. Over time, we found the age profiles of immigration and emigration have been relatively stable, except for persons born in Northeast Asia and in India, whose age patterns both shifted from the *primarily students* classification during the period of 2006–2011 to the *mixed students and young labour* classification during the period of 2011–2016. Notable changes in age profiles were also observed for migration of persons born in the United Kingdom and in North Africa and Middle East. Similarly, we found little differences in the immigration and emigration patterns by sex. Immigrants born in North America and in the Philippines were the only two groups that exhibited substantial differences in their age profiles.

In contrast to changes in age profiles over time and by sex, variations across states and territories in Australia were highly discernible. For example, during the 2011–2016 period, the overall United Kingdom-born immigration flow was classified as *primarily labour*, yet the age profile observed for the corresponding immigration flow to South Australia resembled more the *mixed labour and family* classification. Similarly, the national age profile of India-born immigration was *mixed students and young labour* classification, but the age profile observed in South Australia resembles the *mixed labour and family* classification.

10.5 Conclusion

In this chapter, we have examined age-specific patterns of immigration and emigration and formed a typology for better understanding the primary motivations for migrating from and to Australia. We find that the types of visas used to enter the country may explain many of the differences in the observed age profiles of migration to and from Australia. The main motivations of immigration to Australia appear to be education and labour, with some immigrants bringing their family members with them. The corresponding flows of emigration exhibit lags in the age profiles—implying that once the study or labour is completed, migrants return to their origin country or seek opportunities in other countries.

For our typology of age-specific migration, we gathered detailed immigration and emigration flow data for 19 different birthplace populations by age and sex from 1981 to 2016. However, there were two limitations in these data. First, as we were unable to directly compare the age profiles of migration with the underlying visa entries data, we assumed some relationships between the data sources. Ideally, we would link the two data sets together and examine those age profiles. Second, our data was affected by sparseness especially in the earlier periods of migration and for some states or territories. This limited our analysis to the large states and large flows of migration.

A typology of age-specific migration is useful for simplifying the complexity of migration and for relating migration to life course processes. We have followed Plane and Heins' (2003) framework and analysed origin-destination flows to develop the typology. The origins in our research, however, represented the birthplace of the immigrants and emigrants, which is slightly different from place-to-place migration, as people with different birthplaces may not be migrating from their country or region of birth. Typologies of age-specific migration are also useful for informing population projections. For example, De Beer (2008) proposes the use of argument-based projection models for immigration and emigration to and from the Netherlands. Willekens (2018), furthermore, argues for a causal approach to migration forecasting. While this chapter has not dealt with forecasting migration, we believe the incorporation of a typology of age-specific migration into population forecasting models would inform both argument-based projections and causal-based forecasting models.

In the future, we plan to use this research to study the sources and implications of immigrant population change in Australia. The ages of entry and exit have many implications for the labour force, education, family formation and retirement. Understanding who is coming into the country, and at what ages, is essential for studying the short-term and long-term effects of migration. We hope this research will inspire others to articulate age patterns of migration so that a better understanding of the mechanisms underlying the patterns may occur.

Appendices

Appendix 1 Visa Statistics by Country of Citizenship and Relative Shares of Each Visa Type, 2006–2011

	Student visas				Temporary skilled visas				
Citizenship	Primary applicant	Secondary applicant	Temporary graduate visas	Working holiday visas	Aged under 15	Aged 15–29	Aged 30 and above	Permanent family visas	Permanent skilled visas
Australia	–	–	–	–	–	–	–	–	–
New Zealand	–	–	–	–	–	–	–	–	–
Other Oceania	8551 (29.94%)	4716 (16.51%)	117 (0.41%)	– –	1132 (3.96%)	763 (2.67%)	1523 (5.33%)	6368 (22.30%)	5391 (18.88%)
United Kingdom	10,525 (4.24%)	3612 (1.45%)	602 (0.24%)	36,301 (14.62%)	11,771 (4.74%)	14,316 (5.77%)	30,373 (12.23%)	33,111 (13.33%)	107,711 (43.38%)
NW Europe	38,010 (21.05%)	5718 (3.17%)	707 (0.39%)	63,906 (35.39%)	4875 (2.70%)	12,285 (6.80%)	16,229 (8.99%)	14,651 (8.11%)	24,203 (13.40%)
SE Europe	24,922 (32.44%)	7748 (10.09%)	929 (1.21%)	5928 (7.72%)	1992 (2.59%)	3408 (4.44%)	6605 (8.60%)	13,227 (17.22%)	12,065 (15.70%)
N. Africa and M. East	37,921 (38.98%)	21,532 (22.13%)	684 (0.70%)	536 (0.55%)	1384 (1.42%)	1494 (1.54%)	2924 (3.01%)	16,717 (17.18%)	14,095 (14.49%)
Vietnam	32,078 (57.62%)	2836 (5.09%)	501 (0.90%)	– –	691 (1.24%)	507 (0.91%)	721 (1.30%)	14,724 (26.45%)	3611 (6.49%)
Philippines	9688 (12.27%)	3361 (4.26%)	432 (0.55%)	– –	5937 (7.52%)	3489 (4.42%)	12,893 (16.32%)	15,669 (19.84%)	27,512 (34.83%)
Malaysia	45,362 (55.78%)	4730 (5.82%)	1645 (2.02%)	75 (0.09%)	915 (1.13%)	1561 (1.92%)	2017 (2.48%)	4443 (5.46%)	20,581 (25.31%)
Indonesia	32,931 (56.70%)	7046 (12.13%)	1669 (2.87%)	52 (0.09%)	576 (0.99%)	645 (1.11%)	1445 (2.49%)	6400 (11.02%)	7312 (12.59%)

(continued)

	Student visas				Temporary skilled visas				
Citizenship	Primary applicant	Secondary applicant	Temporary graduate visas	Working holiday visas	Aged under 15	Aged 15–29	Aged 30 and above	Permanent family visas	Permanent skilled visas
SE Asia	60,190 (58.35%)	7980 (7.74%)	1575 (1.53%)	311 (0.30%)	911 (0.88%)	1033 (1.00%)	2399 (2.33%)	16,962 (16.44%)	11,785 (11.43%)
China	234,946 (60.56%)	8875 (2.29%)	13,380 (3.45%)	– –	2883 (0.74%)	3241 (0.84%)	6813 (1.76%)	39,364 (10.15%)	78,485 (20.23%)
NE Asia	105,662 (45.45%)	14,326 (6.16%)	4465 (1.92%)	53,895 (23.18%)	2347 (1.01%)	1713 (0.74%)	5530 (2.38%)	15,150 (6.52%)	29,413 (12.65%)
India	163,970 (42.91%)	41,718 (10.92%)	23,304 (6.10%)	– –	7385 (1.93%)	15,785 (4.13%)	17,239 (4.51%)	22,133 (5.79%)	90,624 (23.71%)
SC Asia	83,440 (47.55%)	23,391 (13.33%)	8160 (4.65%)	– –	1814 (1.03%)	1448 (0.83%)	3181 (1.81%)	16,272 (9.27%)	37,763 (21.52%)
North America	21,627 (25.90%)	2702 (3.24%)	383 (0.46%)	11,943 (14.30%)	4466 (5.35%)	4684 (5.61%)	13,971 (16.73%)	14,191 (16.99%)	9541 (11.43%)
South America	38,087 (52.81%)	9771 (13.55%)	1070 (1.48%)	511 (0.71%)	1574 (2.18%)	2166 (3.00%)	3450 (4.78%)	6781 (9.40%)	8711 (12.08%)
Sub-Saharan Africa	27,957 (21.33%)	10,485 (8.00%)	1693 (1.29%)	– –	8663 (6.61%)	4344 (3.31%)	12,045 (9.19%)	13,291 (10.14%)	52,568 (40.11%)

Notes: 1. Student visas do not include the independent ELICOS sector and non-award sector
2. Working holiday visas only include 20% of the total visas granted
3. Statistics on temporary graduate visas does not include grants made in the financial year of 2006–2007 given data unavailability
Sources: calculations based on visa statistics obtained from the datasets provided by the Department of Home Affairs, available at https://data.gov.au/organization/immi

Appendix 2 Visa Statistics by Country of Citizenship and Relative Shares of Each Visa Type, 2011–2016

Citizenship	Student visas		Temporary graduate visas	Working holiday visas	Temporary Skilled Visas			Permanent family visas	Permanent skilled visas
	Primary applicant	Secondary applicant			Aged under 15	Aged 15–29	Aged 30 and above		
Australia	–	–	–	–	–	–	–	–	–
New Zealand	–	–	–	–	–	–	–	–	–
Other Oceania	11,216 (38.14%)	4562 (15.51%)	610 (2.07%)	– –	1406 (4.78%)	907 (3.08%)	1927 (6.55%)	4936 (16.79%)	3843 (13.07%)
United Kingdom	10,728 (4.01%)	2134 (0.80%)	1511 (0.57%)	44,000 (16.47%)	19,934 (7.46%)	32,959 (12.33%)	45,960 (17.20%)	26,186 (9.80%)	83,816 (31.36%)
NW Europe	36,242 (14.51%)	5376 (2.15%)	1766 (0.71%)	78,083 (31.27%)	11,339 (4.54%)	32,164 (12.88%)	30,588 (12.25%)	15,357 (6.15%)	38,823 (15.55%)
SE Europe	42,076 (29.36%)	14,375 (10.33%)	2972 (2.07%)	16,101 (11.24%)	5745 (4.01%)	10,427 (7.28%)	17,738 (12.38%)	15,241 (10.64%)	18,612 (12.99%)
N. Africa and M. East	37,593 (31.86%)	27,578 (23.37%)	2325 (1.97%)	200 (0.17%)	3048 (2.58%)	2994 (2.54%)	5773 (4.89%)	15,965 (13.53%)	22,533 (19.09%)
Vietnam	45,164 (52.07%)	6426 (7.41%)	4463 (5.14%)	– –	1346 (1.55%)	1853 (2.14%)	1816 (2.09%)	18,088 (20.85%)	7589 (8.75%)
Philippines	19,590 (16.09%)	7130 (5.86%)	3708 (3.05%)	– –	9150 (7.51%)	7340 (6.03%)	17,244 (14.16%)	19,167 (15.74%)	38,438 (31.57%)
Malaysia	43,873 (52.19%)	4283 (5.09%)	5394 (6.42%)	126 (0.15%)	1404 (1.67%)	2810 (3.34%)	3102 (3.69%)	5076 (6.04%)	18,002 (21.41%)
Indonesia	34,626 (55.25%)	7754 (12.37%)	4054 (6.47%)	371 (0.59%)	1138 (1.82%)	1469 (2.34%)	2136 (3.41%)	5656 (9.02%)	5471 (8.73%)
SE Asia	59,469 (52.97%)	11,322 (10.08%)	3213 (2.86%)	457 (0.41%)	1389 (1.24%)	2853 (2.54%)	4711 (4.20%)	17,798 (15.85%)	11,060 (9.85%)

(continued)

Citizenship	Student visas		Temporary graduate visas	Working holiday visas	Temporary Skilled Visas			Permanent family visas	Permanent skilled visas
	Primary applicant	Secondary applicant			Aged under 15	Aged 15–29	Aged 30 and above		
China	283,093 (56.65%)	11,197 (2.24%)	37,657 (7.54%)	1000 (0.20%)	5392 (1.08%)	13,041 (2.61%)	12,025 (2.41%)	52,909 (10.59%)	83,414 (16.69%)
NE Asia	101,188 (38.02%)	16,058 (6.03%)	8053 (3.03%)	75,785 (28.47%)	5679 (2.13%)	5896 (2.22%)	13,751 (5.17%)	16,111 (6.05%)	23,640 (8.88%)
India	114,285 (22.88%)	37,450 (7.50%)	42,074 (8.42%)	– –	25,973 (5.20%)	50,814 (10.17%)	45,814 (9.17%)	31,134 (6.23%)	151,886 (30.41%)
SC Asia	86,815 (32.82%)	33,796 (12.78%)	28,149 (10.64%)	63 (0.02%)	5302 (2.00%)	9390 (3.55%)	8428 (3.19%)	21,732 (8.22%)	70,817 (26.77%)
North America	19,562 (17.69%)	2911 (2.63%)	1327 (1.20%)	15,241 (13.78%)	9086 (8.22%)	10,243 (9.26%)	24,428 (22.09%)	14,507 (13.12%)	13,255 (11.99%)
South America	47,552 (46.95%)	12,421 (12.26%)	3616 (3.57%)	1630 (1.61%)	3153 (3.11%)	4792 (4.73%)	8558 (8.45%)	7581 (7.49%)	11,974 (11.82%)
Sub-Saharan Africa	29,859 (25.20%)	11,505 (9.71%)	5322 (4.49%)	– –	6831 (5.76%)	4727 (3.99%)	10,716 (9.04%)	13,484 (11.38%)	36,051 (30.42%)

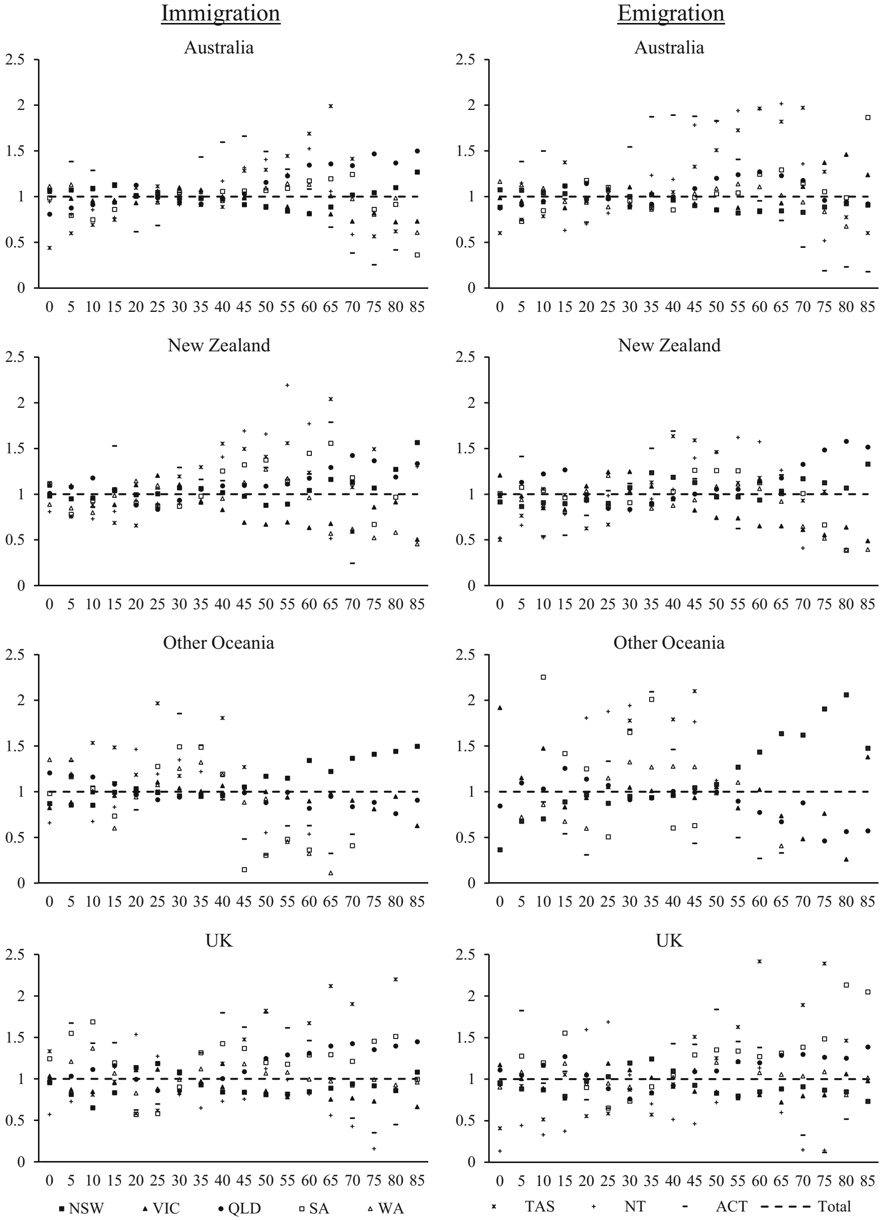
Notes: 1. Student visas do not include the independent ELICOS sector and non-award sector
2. Working holiday visas only include 20% of the total visas granted
Sources: calculations based on visa statistics obtained from the datasets provided by the Department of Home Affairs, available at https://data.gov.au/organization/immi

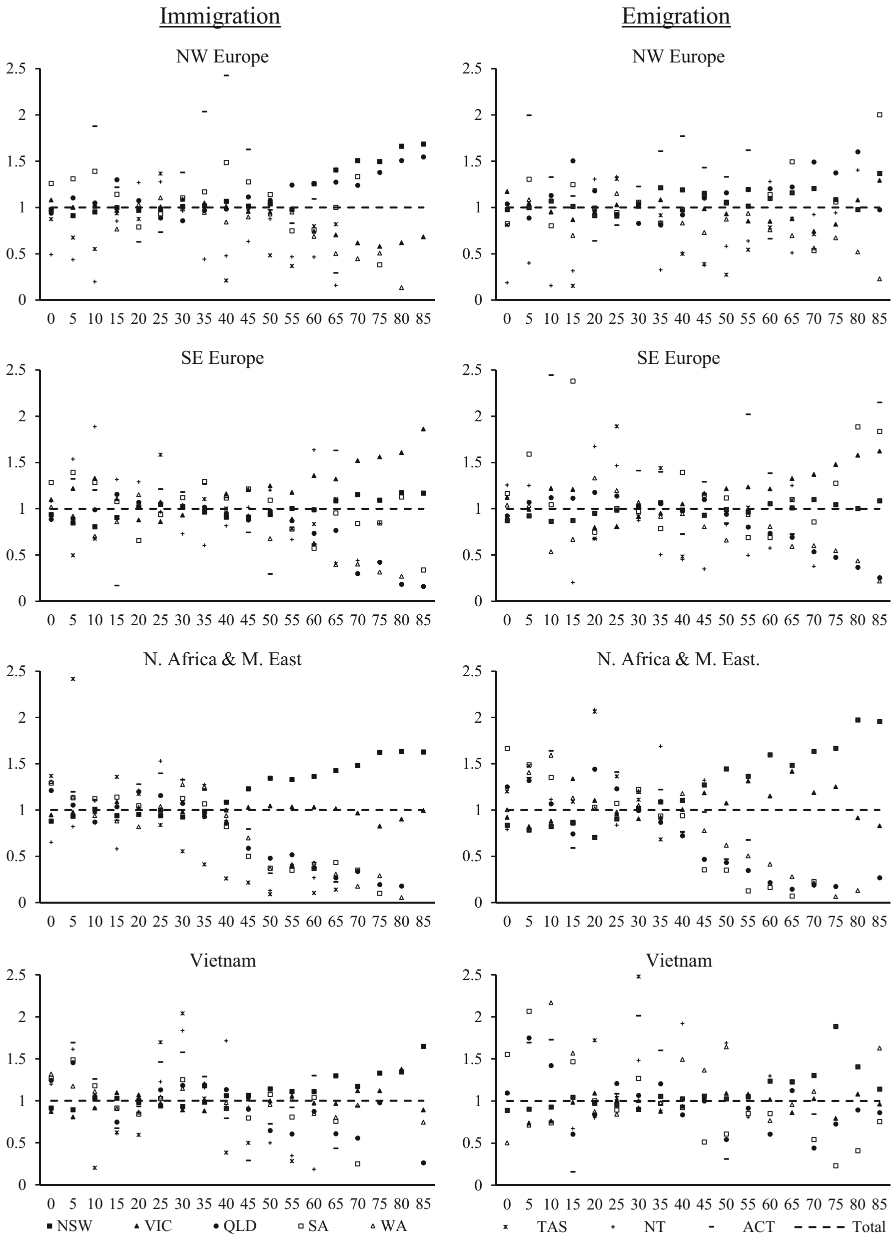
Appendix 3 Visa Statistics by Sex for the Philippines and North America Citizenships and Relative Shares of Each Visa Type, 2011–2016

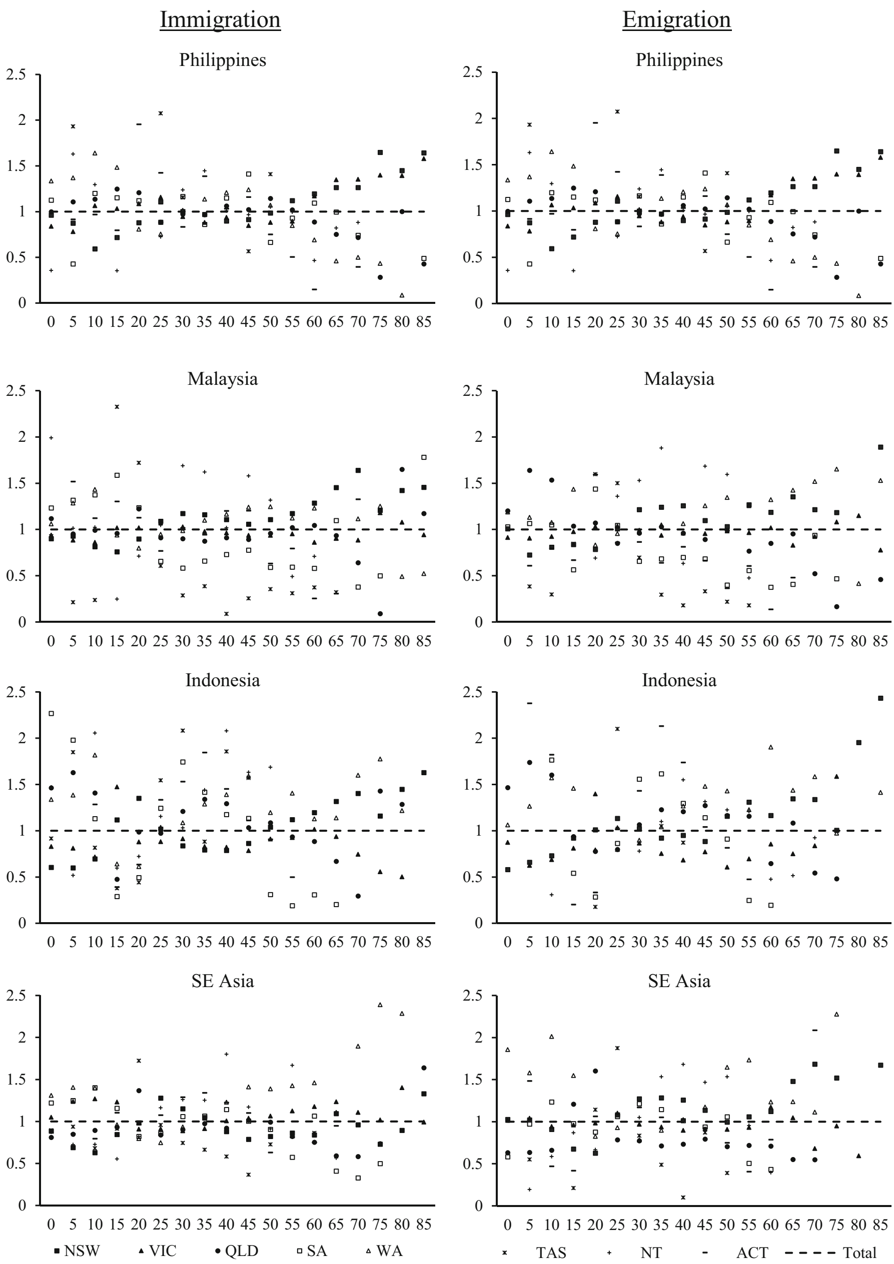
Citizenship	Student Visas		Temporary graduate visa	Working holiday visas	Temporary skilled visas			Permanent family visas	Permanent skilled visas
	Primary applicant	Secondary applicant			aged under 15	aged 15–29	aged 30 and above		
Philippines									
Male	6702 (11.90%)	4324 (7.68%)	1530 (2.72%)	– –	4774 (8.48%)	3711 (6.59%)	11,163 (19.82%)	4816 (8.55%)	19,292 (34.26%)
Female	12,888 (19.69%)	2806 (4.29%)	2178 (3.33%)	– –	4371 (6.68%)	3629 (5.54%)	6081 (9.29%)	14,351 (21.93%)	19,146 (29.25%)
North America									
Male	8440 (15.66%)	1412 (2.62%)	559 (1.04%)	6324 (11.74%)	4588 (8.51%)	4924 (9.14%)	15,108 (28.03%)	5824 (10.84%)	6695 (12.42%)
Female	11,117 (19.62%)	1499 (2.65%)	763 (1.35%)	8916 (15.74%)	4498 (7.94%)	5319 (9.39%)	9315 (16.44%)	8665 (15.30%)	6560 (11.58%)

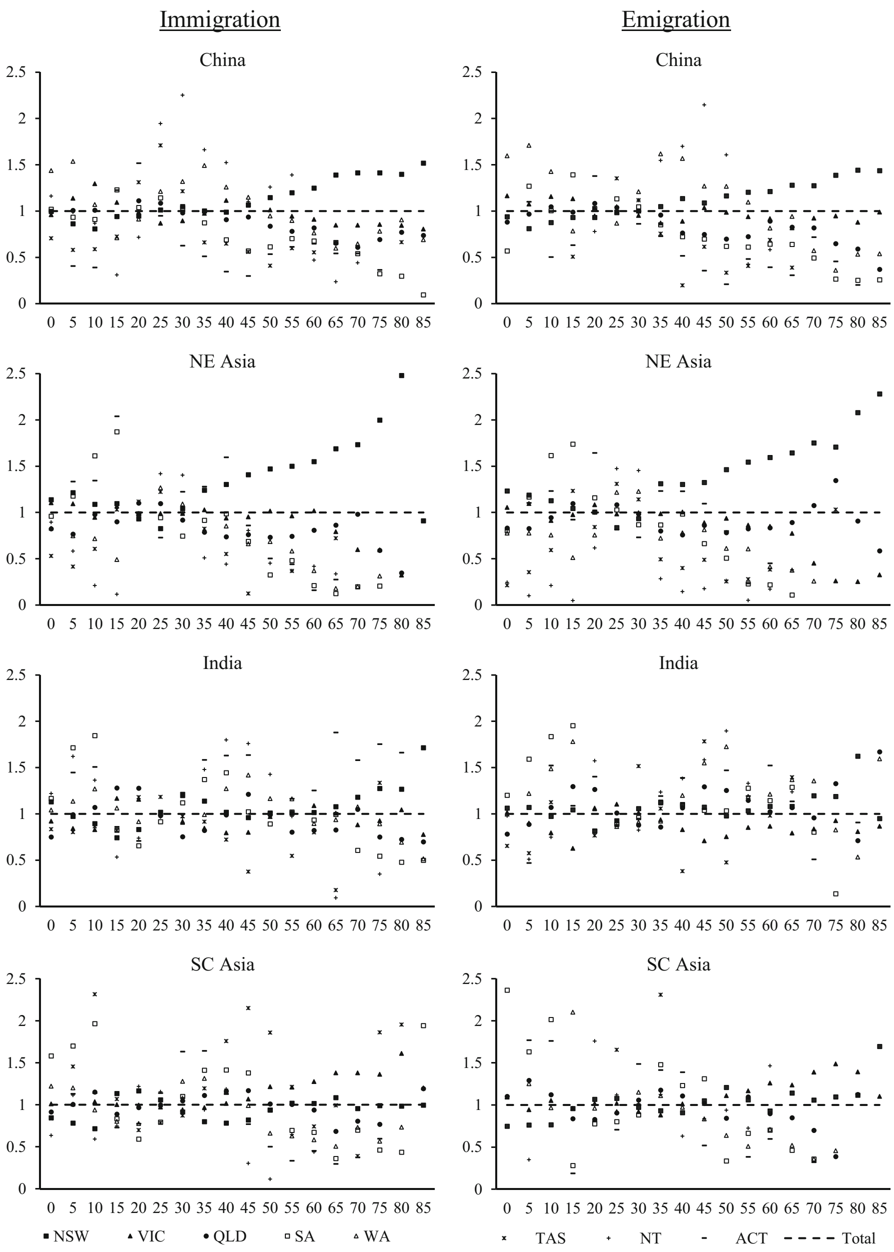
Notes: 1. Student visas do not include the independent ELICOS sector and non-award sector
2. Working holiday visas only include 20% of the total visas granted
Sources: calculations based on visa statistics obtained from the datasets provided by the Department of Home Affairs, available at https://data.gov.au/organization/immi

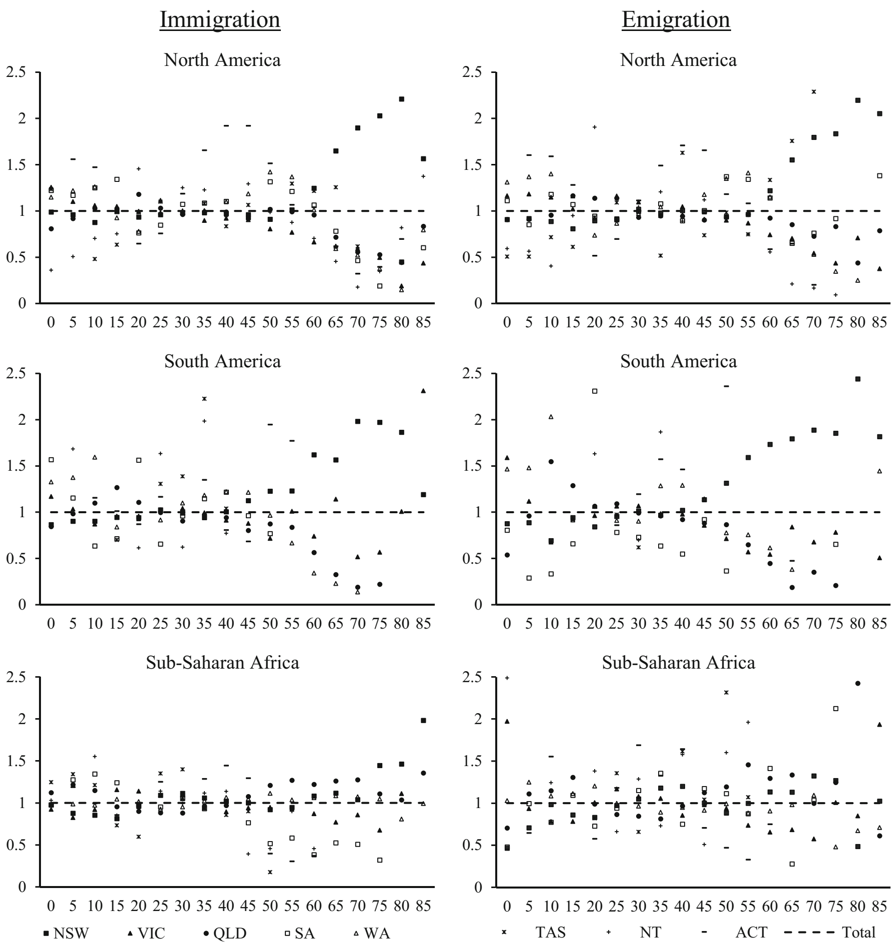
Appendix 4 Age-Proportion Ratios (State/Total) by Birthplace, 2011–2016




Immigration
Emigration
NW Europe
SE Europe
N. Africa & M. East
N. Africa & M. East.
Vietnam
2.5
2
1.5
1
0.5
0
0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85
NSW
VIC
QLD
SA
WA
TAS
NT
ACT
Total


Immigration
Emigration
Philippines
Philippines
Malaysia
Malaysia
Indonesia
Indonesia
SE Asia
SE Asia
0 0.5 1 1.5 2 2.5
0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85
NSW
VIC
QLD
SA
WA
TAS
NT
ACT
Total


Immigration
Emigration
China
China
NE Asia
NE Asia
India
India
SC Asia
SC Asia
2.5
2
1.5
1
0.5
0
0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85
NSW
VIC
QLD
SA
WA
TAS
NT
ACT
Total



Note: "Total" denotes the national-level age proportion of each birthplace-specific migrant group

References

ABS (2018) Migration, Australia 2016–17. Cat. no. 3412.0, Australian Bureau of Statistics, Canberra. Retrieved from: http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/3412.0Main+Features32016-17?OpenDocument

Bernard A, Bell M (2015) Smoothing internal migration age profiles for comparative research. Demogr Res 32(33):915–948

Bernard A, Bell M, Charles-Edwards E (2014) Life-course transitions and the age profile of internal migration. Popul Dev Rev 40(2):213–239

Castro LJ, Rogers A (1983) What the age composition of migrants can tell us. Popul Bull U N 15:63–79

De Beer J (2008) Forecasting international migration: time series projections vs argument-based projections. In: Raymer J, Willekens F (eds) International migration in Europe: data, models and estimates. Wiley, Chichester, pp 283–306
DIBP (2013) Australia's Migration Trends 2011–12. Department of Immigration and Border Protection, Canberra
DIBP (2015) Australia's Migration Trends 2013–14. Department of Immigration and Border Protection, Canberra
DIBP (2016a) Australia's Migration Trends 2015–16. Department of Immigration and Border Protection, Canberra
DIBP (2016b) Student Visa and Temporary Graduate Visa Programme Trends 2009–10 to 2015–16. Department of Immigration and Border Protection, Canberra
DIBP (2016c) Visitor Visa Programme Bi-annual Report Ending at 30 June 2016. Department of Immigration and Border Protection, Canberra
DIBP (2016d) Working Holiday Maker Visa Programme Report 30 June 2016. Department of Immigration and Border Protection, Canberra
Home Affairs (2018a) 2012–13 Permanent Additions. Department of Home Affairs, Canberra. Retrieved from https://data.gov.au/dataset/permanent-additions-to-australia-s-resident-population/resource/04a427e5-aa32-4e9e-8890-ee851aad4f12
Home Affairs (2018b) 2014–15 Permanent Additions. Department of Home Affairs, Canberra. Retrieved from https://data.gov.au/dataset/permanent-additions-to-australia-s-resident-population/resource/17a7ebab-2895-41cd-b7e9-d8de6367999e?inner_span=True
Hugo G (2008) Immigrant settlement outside of Australia's capital cities. Popul Space Place 14:553–571
Hugo G (2009) Flows of immigrants 1993–2008: Australia. In: Higley J, Nieuwenhuysen J, Neerup S (eds) Nations of immigrants: Australia and the USA compared. Edward Elgar Publishing, Cheltenham, pp 22–41
Hugo G, Harris K (2011) Population distribution effects of migration in Australia. Department of Immigration and Citizenship, Canberra. Retrieved from https://www.border.gov.au/ReportsandPublications/Documents/research/full-report.pdf
Jupp J (2001) The Australian people: an encyclopaedia of the nation, its people and their origins. Cambridge University Press, Cambridge
Khoo S-E (2003) A greater diversity of origins. In: Khoo S-E, McDonald P (eds) The transformation of Australia's population. University of New South Wales Press, Sydney, pp 158–184
Markus A, Jupp J, McDonald P (2009) Australia's immigration revolution. Allen & Unwin, Crows Nest
Parliament of Australia (2010) Overseas students: immigration policy changes 1997 – May 2010. Parliament of Australia, Canberra. Retrieved from https://www.aph.gov.au/About_Parliament/Parliamentary_Departments/Parliamentary_Library/pubs/BN/0910/OverseasStudents
Pittenger DB (1974) A typology of age-specific net migration rate distributions. J Am Inst Plann 40(4):278–283
Pittenger DB (1978) On making flexible projections of age-specific net migration. Environ Plan A 10:1253–1272
Plane DA (1993) Demographic influences on migration. Reg Stud 27(4):375–383
Plane DA, Heins F (2003) Age articulation of U.S. inter-metropolitan migration flows. Ann Reg Sci 37:107–130
Plane DA, Henrie CJ, Perry MJ (2005) Migration up and down the urban hierarchy and across the life course. Proc Natl Acad Sci 102(43):15313–15318
Raymer J, Baffour B (2018) Subsequent migration of immigrants within Australia, 1981–2016. Popul Res Policy Rev 37(6):1053–1077
Raymer J, Rogers A (2008) Applying model migration schedules to represent age-specific migration flows. In: Raymer J, Willekens F (eds) International migration in Europe: data, models and estimates. Wiley, Chichester, pp 175–192

Raymer J, Shi Y, Guan Q, Baffour B, Wilson T (2018) The sources and diversity of immigrant population change in Australia, 1981–2011. *Demography* 55(5):1777–1802

Raymer J, Bai X, Liu N, Wilson T (forthcoming) Estimating a consistent and detailed time series of immigration and emigration for sub-state regions of Australia. *Applied Spatial Analysis and Policy*. https://doi.org/10.1007/s12061-019-09310-w

Richards E (2008) Destination Australia. University of New South Wales Press, Sydney

Rogers A, Castro LJ (1981) Model migration schedules. IIASA Research Report RR-81-030. International Institute for Applied Systems Analysis, Laxenburg

Rogers A, Racquillet R, Castro LJ (1978) Model migration schedules and their applications. Environ Plan A 10:475–502

Rogers A, Little J, Raymer J (2010) The indirect estimation of migration: methods for dealing with irregular, inadequate, and missing data. Springer, Dordrecht

Willekens F (1999) The life course: models and analyses. In: Van Wissen LJG, Dykstra P (eds) Population issues: an interdisciplinary focus. Kluwer Academic/Plenum Publishers, New York, pp 23–51

Willekens F (2018) Towards causal forecasting of international migration. Vienna Yearb Popul Res 16:1–20

Wilson T, Raymer J (2017) The changing shape of Australia's overseas-born population. Popul Soc 545:1–4

Chapter 11
Modelling Inter-urban Migration in an Open Population Setting: The Case of New Zealand

Omoniyi B. Alimi, David C. Maré, and Jacques Poot

Abstract In this chapter, we revisit the modelling of gross inter-urban migration flows in New Zealand. As in previous work, we identify a range of geographic, demographic, economic and climatic characteristics of urban areas, which are statistically significant determinants of migration. However, we argue that in a small but open population such as New Zealand (in which one quarter of the resident population is foreign born and one sixth of the New Zealand-born population lives abroad), inter-urban migration should be modelled jointly with rural-urban and international migration. We proceed to estimate a modified gravity model of migration in which the flow matrix is augmented with rural-urban and international migration. Migration data are obtained from four successive population censuses since 1996. We find notable differences in the impact of migration determinants when comparing urban-urban, urban-rural and urban-world migration flows. The estimation of these models is straightforward and does not require collection of data on rural areas or foreign countries. Hence, the method can be easily applied to other case studies in which international and/or rural-urban migration are important components of population churn.

Keywords Gravity model · Internal migration · International migration · New Zealand

O. B. Alimi
Waikato Management School, University of Waikato, Hamilton, New Zealand
e-mail: niyia@waikato.ac.nz

D. C. Maré
Motu Economic and Public Policy Research, Wellington, New Zealand
e-mail: dave.mare@motu.org.nz

J. Poot (✉)
National Institute of Demographic and Economic Analysis (NIDEA), University of Waikato, Hamilton, New Zealand

Department of Spatial Economics, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
e-mail: h.j.poot@vu.nl

R. S. Franklin (ed.), *Population, Place, and Spatial Interaction*, New Frontiers in Regional Science: Asian Perspectives 40,
https://doi.org/10.1007/978-981-13-9231-3_11

11.1 Introduction

In our increasingly urbanised world, subnational and cross-border migration of people between cities is a prominent aspect of demographic change. In the developed world, where sub-replacement fertility is commonplace, inter-urban migration has become the main driver of population change—a reality already predicted by Zelinsky (1971) as the fifth and final stage of the mobility transition. At the same time, the world has witnessed for several decades a divergence in the fortunes of cities—with the largest services-driven cities being triumphant (Glaeser 2012) and with many small- and medium-sized manufacturing-driven cities declining. This regional divergence is pronounced in the European Union (e.g. Iammarino et al. 2019), but it can also be observed in other parts of the developed world, including in New Zealand (Alimi et al. 2016). It is no exaggeration to say that the rise of populist politicians in many countries and the vote for Brexit in the UK are expressions of anger by those in declining cities and regions.

Are people also voting with their feet, through migrating to the largest cities (that are likely to offer higher wages, more employment opportunities and attractive amenities) in their own country or even across the border? This is the issue that we revisit in the present chapter by means of data for New Zealand. New Zealand is of interest, because this country is highly urbanised (with four out of five people living in urban areas) and has also one of the highest rates of internal and cross-border population mobility in the world.

Although comparison of migration rates across countries is difficult, due to differences in definitions of who constitutes a migrant, differences in measurement methods and the diversity in the size and shape of spatial areas over which migration is measured across countries (Greenwood 1997), studies that have attempted to compare migration rates across countries have found New Zealanders to be highly mobile (Long 1991; Esipova et al. 2013). Around half of all New Zealanders change their residence at least once during a 5-year period (Maré et al. 2007). For an open population, such as New Zealand, it is important to consider internal and cross-border flows simultaneously (King and Skeldon 2010). The 2013 population census of New Zealand showed that nationally around one quarter of the population are foreign born. In major cities like Auckland, the foreign-born account for almost two in five people (Statistics New Zealand 2013). At the same time, New Zealand has a significant diaspora. It is estimated that about one in six of the New Zealand-born population lives abroad.[1]

[1]According to the census, the usually resident population of New Zealand in 2013 was 4.2 million, of whom 75% were New Zealand born. At the time of the 2016 Australian census, there were 518,000 New Zealand-born residents of Australia. The total New Zealand diaspora is, in terms of citizenship, estimated to be around 750,000 (https://en.wikipedia.org/wiki/New_Zealanders), but less in terms of country of birth—perhaps around 600,000. Hence, of the global number of New Zealand born that equals $0.75 \times 4{,}200{,}000 + 600{,}000 = 3{,}750{,}000$, about one in six lives abroad.

Empirical research has shown extensively that migration is not random with respect to people and places. Migrants are self-selected in terms of a range of personal and locational characteristics. As Plane (1993) shows, demographic characteristics of people and places are fundamental drivers of migration. The strongest predictor is undoubtedly a person's age (see e.g. Plane and Heins 2003; Hunt and Mueller 2004), with young people aged 18 to 29 having the highest geographic mobility rates. However, there is also plenty of evidence in the literature regarding the importance of other personal characteristics, such as marital status, education, employment status and networks.[2] Differences between places in terms of income, regional unemployment, housing and amenities (such as climate) matter too.[3] Furthermore, patterns of migration change over time, and there is heterogeneity in behaviour across population cohorts (Plane 1993). Finally, given the growth in the percentage of foreign born in the developed world, authors have been increasingly paying attention to the interaction between internal and international migration in population redistribution.[4]

The dominant methodological framework for modelling gross migration flows, originating as far back as Ravenstein (1885, 1889), remains the gravity model in which migration between places is positively related to the scale of population and economic activity in origin and destination, but negatively related to the distance between them. Social scientists have witnessed a resurgence of the gravity model in recent years due to its application to international migration flows and the growing attention to incorporating spatial spillovers or systemic effects in gravity models (Poot et al. 2016).

We use in this chapter a modified gravity model to examine gross internal migration flows between the main and secondary urban areas in New Zealand during four periods (1991–1996, 1996–2001, 2001–2006 and 2008–2013), while simultaneously accounting for flows between rural and urban areas, as well as cross-border flows. The way in which we implement this estimation approach is straightforward and can be also easily applied in other contexts with open populations (i.e. populations in which both internal and cross-border flows are important contributors to regional population change). While there has already been much research on the socio-economic determinants of internal migration in New Zealand,[5] the present study examines some additional factors that have not received attention previously. Specifically, we consider the impact of personal income inequality and

[2]Regarding such personal characteristics, see, for example, on marital status, Maxwell (1988); on education, Greenwood (1969) and Caldwell (1968); on employment status, Pissarides and McMaster (1990); and on networks, Curran and Rivero-Fuentes (2003), Pedersen et al. (2008) and Michaelides (2011).

[3]On income, see, for example: Sjaastad (1962) and Greenwood (1969); on regional unemployment, Westerlund (1998) and Cebula and Alexander (2006); on housing, Jeanty et al. (2010) and Modestino and Dennett (2013); and on climate, Graves (1979); Rappaport (2007) and Poston et al. (2009).

[4]See, for example, Frey (1996), Borjas et al. (1996), Card and DiNardo (2000), Borjas (2006) and Glitz (2012).

[5]See, for example, Poot (1986), Maré and Timmins (2000), Maré and Choy (2001), Maré et al. (2007), Sloan (2013), Poot et al. (2016) and Cameron and Poot (2019).

the subnational distribution of international migrants. We also update evidence on socio-economic determinants of internal migration examined in earlier studies.

As in previous work, we find that the basic gravity model provides a remarkably good statistical description of New Zealand's migration flows. We also find that real income growth deters outward migration and attracts inward migration. Additionally, growing income inequality leads to more gross migration. The impacts of these income variables differ between internal migration and cross-border migration. We also investigate the so-called 'skating rink' hypothesis that suggests that net inward international immigration leads to local internal outward migration. In contrast to this hypothesis, but similar to what Card and DiNardo (2000) found in the USA, we find that international migrants and inward migrants appear complements rather than substitutes: urban areas with greater net inward international migration have also greater inward internal migration and less outward internal migration. At the same time, the areas which have the largest proportions of foreign born are the areas with relatively lower geographic mobility. With respect to personal characteristics, we find that homeownership lowers geographical mobility, but youthfulness increases it. Migrants are also disproportionally attracted to urban areas with high levels of human capital. In terms of climate variables, wetter places trigger less internal migration, but this is not the case for cross-border flows. More sunshine discourages outward migration—at least internally. Finally, in line with Poot (1986), we find that spatial interaction in the form of internal migration is relatively more intense between the six major urban centres of New Zealand than between pairs of other cities.

The next section outlines the methodology we use. This is followed by a description of the New Zealand data and sources. The penultimate section reports the results. Some final comments are provided in the concluding section.

11.2 Methodology

The basic gravity model of migration relates the number of people migrating between two places to the population of both places and the distance between them:

$$M_{ij} = \alpha \frac{(P_i)^{\beta_1} \left(P_j\right)^{\beta_2}}{\left(D_{ij}\right)^{\delta}} \tag{11.1}$$

where M_{ij} is the migration flow between origin city i and destination city $j(i, j = 1,2,\ldots,R)$, P_i is population at origin, P_j is population at destination and D_{ij} is distance between origin i and destination j, for example, road distance in kilometres between the city centres. The model is typically estimated by log-linearising Eq. 11.1 and using OLS to estimate the parameters α, β_1, β_2 and δ which are interpreted, respectively, as a proportionality constant, the origin population elasticity, destination population elasticity and distance elasticity of migration (with $\gamma = -\delta$ and adding an error term ε_{ij})

$$LnM_{ij} = Ln\alpha + \beta_1 LnP_i + \beta_2 LnP_j + \gamma LnD_{ij} + \varepsilon_{ij} \tag{11.2}$$

Despite this extremely simple model having often a remarkably good statistical fit, several weaknesses of this approach have been identified as far back as the 1970s. Weaknesses include econometric specification issues, a lack of socio-economic theoretical underpinnings and data issues such as dealing with zero flows (see e.g. Ramos 2016 for a review). Additionally, the basic gravity model does not account for spatial spillovers and/or spatial heterogeneity of parameters (see LeSage and Pace 2008, 2009; Peeters 2012). Recently, Cameron and Poot (2019) focused on interpretational issues of regression equations such as Eq. 11.2 and argue that OLS and fixed (origin and destination) effect estimates of the coefficients of the basic gravity model are biased. Additionally, they find that population elasticities obtained from fixed origin and destination effects estimation of Eq. 11.2 represent growth rate rather than level effects.

One enduring criticism of the basic form of the gravity model has been that it suffers from omitted variable bias by failing to incorporate other important people and place characteristics that could determine migration flows. To overcome this bias, researchers incorporate several variables that characterise origins, destinations, the composition of the migration flows and (i, j)-specific facilitators or inhibitors of migration other than distance. The modified gravity model is then of the form:

$$\begin{aligned} LnM_{ij} = {} & Ln\alpha + \beta_1 LnP_i + \beta_2 LnP_j + \gamma LnD_{ij} \\ & + \textstyle\sum_{k=1}^{K} \mu_k X_{ki} + \sum_{l=1}^{L} \nu_l X_{lj} + \sum_{n=1}^{N} \pi_n F_{nij} + \varepsilon_{ij} \end{aligned} \tag{11.3}$$

where X_{ki} and X_{lj} are the additional migration-determining characteristics in origin and destination areas, respectively, and F_{nij} represents a set of variables that measures spatial friction between i and j other than distance (e.g. whether travel between i and j involves a ferry across water, direct flights or a country border).

Although the gravity model has long been a popular workhorse of internal migration research (see Greenwood 1975, 1997; Ramos 2016) and has also been increasingly applied to international migration (e.g. Karemera et al. 2000; Mayda 2010), there have been to date few attempts to use the gravity approach to simultaneously model both internal and international flows (but see e.g. Poot et al. 2016). In this chapter, we use again a modified gravity model to concurrently examine the determinants of gross international and internal migration. Our model differs from the one in Eq. 11.3 in one key aspect: we allow for heterogeneity with respect to internal (inter-urban and rural-urban/urban-rural) flows and international flows. As in Poot et al. (2016), we use New Zealand as a case study. New Zealand provides a good example given that, as noted earlier, New Zealanders have a high rate of internal mobility and a high rate of international migration.

To expand beyond inter-urban flows, we consider the population and other characteristics of the generic overseas origins and destinations to be undefined, because the available data are not disaggregated in terms of specific origins and destinations of international migrants. Similarly, although the total rural population is known, ‘rural’ is a spatially dispersed origin or destination rather than a specific location. Hence we also classify its population, distance and other characteristics as

From / To	Destination		
Origin	UA1 UA2 UA3 ...UA40	Rural	International
UA 1	X		
UA 2	X		Emigrants from UA ($UW = 1$)
UA 3	X		
…….	Inter-urban flows (UU=1)	Urban to rural migrants ($UR = 1$)	
UA40	X		
Rural	Rural to urban migrants ($RU = 1$)	X	Emigration from rural areas (UU=UR=UW=RU=WU=0)
International	Immigrants to UA ($WU = 1$)	Immigrants to rural areas (UU=UR=UW=RU=WU=0)	X

Fig. 11.1 Gross migration matrix and dummy variables defining types of migration flows. (Note: X refers to the main diagonal of the gross migration matrix. These cells represent intra-area mobility which is not observed and thus excluded from the analysis)

unknown. However, our methodology permits us to examine whether population and other variables of urban areas have a different impact on migration when the flows under consideration are inter-urban as compared with rural-urban or abroad-urban. We define five dummy variables:

- $UU_{ij} = 1$ if and only if both the origin i and the destination j are urban areas, and 0 otherwise
- $UR_{ij} = 1$ if and only if the origin i is an urban area and the destination j is rural (i.e. these correspond to the urban to rural flows), and 0 otherwise
- $RU_{ij} = 1$ if and only if the origin i is rural and the destination j is an urban area (i.e. these correspond to rural to urban flows), and 0 otherwise
- $UW_{ij} = 1$ if and only if the origin i is an urban area and the destination j is abroad (i.e. these correspond to the emigration flows), and 0 otherwise
- $WU_{ij} = 1$ if and only if the origin i is abroad and the destination j is an urban area (i.e. these correspond to the immigration flows), and 0 otherwise.

Figure 11.1 shows the origin-destination matrix and the dummy variables accounting for each type of flow. Through the use of the dummy variables defined above, the values assigned to variables measuring characteristics of rural areas or overseas areas, and the distances between urban areas and abroad, or between urban areas and rural areas, become irrelevant.[6] Note that in this setup, we cannot account for flows between international and rural areas, given that none of the migration model's determinants are defined for these flows.

Consider now a pooling of gross migration matrices for several periods. To simplify notation, let X represent the area attributes that are observed in each period for each urban area. We can now specify a modified gravity model that concurrently examines the effect of these area characteristics on gross inter-urban, rural-urban and international-urban migration flows as follows:

$$\begin{aligned} lnM_{ij,t} = {} & \text{Intercepts} + (\beta_{1UU}UU + \beta_{1UR}UR + \beta_{1UW}UW)LnP_{i,t-1} \\ & + (\beta_{2UU}UU + \beta_{2RU}RU + \beta_{2WU}WU)LnP_{j,t-1} + \gamma LnD_{ij} \\ & + \gamma_{AG} + \gamma_{\text{Island}} + (\mu_{UU}UU + \mu_{UR}UR + \mu_{UW}UW)X_{i,t-1} \\ & + (\nu_{UU}UU + \nu_{RU}RU + \nu_{WU}WU)X_{j,t-1} + \tau_t + \varepsilon_{ij,t} \end{aligned} \tag{11.4}$$

where $M_{ij,t}$, $P_{i,t-1}$, $P_{j,t-1}$, D_{ij} are as defined earlier in Eq. 11.1, τ_t are period dummies and $X_{i,t-1}$ and $X_{j,t-1}$ are the economic and demographic attributes of urban areas lagged one period.[7] Apart from economic and demographic attributes, we also consider climatic variables measured by average rainfall and sunshine. We include two additional dummy variables to capture some New Zealand-specific geographic effects. We include an agglomeration-interaction dummy (γ_{AG}) to capture flows between the six urban areas that make up New Zealand's six largest cities. We also include an interisland dummy (γ_{Island}) to capture the barrier that the Cook Strait forms between the North Island and the South Island.

An important issue with the specification of gravity models in logarithms is dealing with cases of migration flows between specific origins and destinations being zero.[8] Here, we set $M_{ij} = 0.5$, where the reported migration flow is 0.[9] Zero

[6]In the estimation in Stata, we have set these values to 0.

[7]The economic and demographic attributes of urban areas are lagged one period to limit the impact of potential endogeneity of the regressors. It is expected that economic and demographic conditions at the time of the previous census are likely to have impacted the migration flow over the subsequent intercensal period. At the same time, a shock to migration impacts on current and future population, but not on past population. Lagging the explanatory variables is even more appropriate when we consider that the migration data are arrived at by checking whose area of residence is different from where they were 5 years ago. Given that expectations are often based on extrapolations from the past, a person whose current residence in the 2013 census is reported as Wellington but who resided in Auckland at the time of the 2006 census is more likely to have changed residence because of conditions in Auckland and Wellington prior to 2006, and including 2006, than subsequently.

[8]Because migration flows in gravity models are specified in logarithms, zero flows present a challenge as the logarithm of zero is not defined. See, e.g. Ramos (2016) for alternative approaches.

[9]To preserve confidentiality, New Zealand census counts are rounded to multiples of 3: an actual count of 0 is reported as such, but an actual count of 1 is rounded down to 0 with probability 2/3 and rounded up to 3 with probability 1/3, with the reverse probabilities for rounding a count of 2. If the low frequencies were uniformly distributed, a rounded value of 0 is therefore reported in 2/3 of the cases rather than 1/3. However, the distribution of low frequencies is unlikely to be uniform, with 0 likely to be much more common than 1 or 2, particularly in migration matrices referring to small areas or relatively small groups.

migration flows in our data form only around 5% of the total flows in all census periods. Most results are not sensitive to the exclusion of zero flows, except for a few instances that are elaborated in the discussion of the results.[10]

11.3 Data

We use data from all five population censuses of population and dwellings between 1991 and 2013. However, our model covers only four migration periods (1991–1996, 1996–2001, 2001–2006, 2008–2013) since our independent variables are specified in lags; hence 1991–1996 migration is linked to 1991 socio-economic conditions. All data are from the census, apart from the distance variable sourced from Google Maps and the climate variables sourced from the National Institute of Water and Atmospheric Research (NIWA). Censuses were held every 5 years, apart from in 2011 when the census was postponed until 2013 due to the Canterbury earthquakes. The dependent variable (M_{ij}) represents the gross migration flows between origin i and destination j for each of the 40 urban areas; or flows to/from rural areas from/to each urban area; or international emigration and immigration from/to each urban area. The census includes information on residents who have newly arrived from overseas (immigration). However, since the census includes only people who are actually in New Zealand at the time of the census, emigration from New Zealand is not recorded in census data, but we estimate this by means of a residual method.[11]

A migrant is defined as someone aged between 25 and 54 who is living in an urban area that is different from where he or she was living at the time of the previous census. The focus is specifically on people aged 25 to 54 to capture economic and labour market determinants of migration and exclude other types of migration that may bias the results, such as moves made by young people in pursuit of tertiary education or by retirees in pursuit of lifestyle options.

Statistics New Zealand 2013 classifications are used to define the urban areas used in this study. Statistics New Zealand considers an urban area to be a region with a population of 1000 or more. Population size is not the only criterion to classify urban areas—factors such as remoteness and location of employment of most of the

[10]Estimations in which zero migration flows are excluded are available from the authors upon request.

[11]Given that censuses are held at the same time every 5 years (but 7 years between the 2006 and 2013 census), cohorts can be followed over time. After accounting for immigration, internal migration and age and sex specific mortality rates, emigration can be calculated as the residual change in the size of a cohort. Of course the resulting numbers are measured with some error, due to census undercounting, etc.

population are also used to further differentiate the type of urban area.[12] The focus of this study is on the 40 main and secondary urban areas. The independent variables are in six categories:

- *Population variable*—Population of 25–54 in each urban area.
- *Distance*—Road travel distance in kilometres between the centres of the two urban areas. Travel distances are estimated using Google Maps.[13]
- *Geographic variables*—An interisland flow dummy to capture the additional barrier to migration represented by Cook Strait between New Zealand's North and South Islands. An agglomeration-interaction dummy captures flows between the urban areas of New Zealand's six largest cities.
- *Economic variables*—Growth rate of real average income in each urban area, changes in income distribution (measured by the Theil index of income inequality), proportion of people living in their own home and proportion of the population with a high level of education (defined as those with at least a bachelor's degree).
- *Demographic variables*—Ratio of the size of the 15 to 24 age group to the 25 to 54 group (i.e. share of the young); ratio of the size of the 55 to 64 age group to the 25 to 54 group (i.e. share of the aged); percentage of the population in an urban area that is foreign born; and an international migration effectiveness ratio which is defined as the ratio of the number of international immigrants to international emigrants in each urban area.
- *Climatic variables*—The 30-year average of sunshine hours in each urban area and the 30-year average of rainfall measured in millimetres.[14]

Summary statistics of these variables are given in Table 11.1. In each census period, there are 1560 (40 × 39) inter-urban flows, 80 (40 × 2) rural-urban flows and 80 (40 × 2) international-urban flows. This leads to a total of 1720 migration flow observations in each census period. When we pool across four censuses, the total number of observations becomes therefore 6880. The distance measure, interisland dummy and agglomeration dummy only apply to inter-urban flows. Hence, for those flows, we have 1560 (40 × 39) observations, making a total of 6240 observations pooled across 4 censuses. There are 40 urban areas, but because we pool across 4 censuses, the number of observations for urban area characteristics equals 160.

[12]Three types of urban areas exist: a main urban area is one with a population of at least 30,000 people, a secondary urban area is one with a population of less than 30,000 people but where more than 20% of the employed population works in a main urban area and a minor urban area is one with a population of less than 30,000 people and where less than 20% of the employed population works in a main urban area.

[13]Manukau city centre was the reference point for the South Auckland urban area, Henderson for West Auckland, North Shore Information centre for North Auckland and Auckland city centre for the Central Auckland urban area.

[14]Average rainfall and sunshine information was available for only 20 out of the 40 urban areas. For urban areas where data were unavailable, data from the nearest urban area within a 100 km range were used to proxy for the missing information.

Table 11.1 Summary statistics of all variables, pooled cross sections from 1996 to 2013

Variables	Obs.	Mean	Std. dev.	Min	Max
Paired observations for all areas pooled across all periods from 1996 to 2013					
Migration flow (includes rural and international flows)	6880	286.16	1452.02	0.50	37,551.00
Paired observations across urban areas pooled across all periods from 1996 to 2013					
Interisland dummy	6240	0.41	0.49	0.00	1.00
Agglomeration dummy	6240	0.05	0.21	0.00	1.00
Travel time (km)	6240	568.00	390.98	10.20	1784.00
Urban area characteristics pooled across all periods from 1996 to 2013					
Resident population	160	31,835	41,990	3732	193,188
Real growth rate in income	160	0.06	0.04	−0.06	0.17
Growth in income inequality	160	0.02	0.09	−0.23	0.25
High qualification rate	160	0.12	0.07	0.04	0.45
Ratio of the population aged 15 to 24 to the population aged 25 to 54 (youthfulness)	160	0.34	0.07	0.23	0.60
Ratio of the population aged 55 to 64 to the population aged 25 to 54 (agedness)	160	0.23	0.05	0.12	0.37
Proportion of the 25–54 population that is foreign born	160	0.17	0.09	0.06	0.48
Proportion of the 25–54 population living in own home	160	0.71	0.10	0.35	0.88
International emigrants	160	2211	3847	0	24,602
International immigrants	160	3048	5817	84	37,551
Time invariant characteristics					
Rainfall (mm)	40	1105.40	376.21	573.00	2875.00
Sunshine (h)	40	2025.38	182.15	1585.00	2409.00

Climate variables are measured as averages over the whole period. Hence, we have only 40 observations on climate.

11.4 Results

We begin this section with the basic gravity model used to illustrate the additional insights that we gain by simultaneously modelling internal and international migration. We then proceed to discuss the results from the modified gravity model.

Table 11.2 presents the result of the basic gravity model with population and distance as the only explanatory variables. The model is run for a pooled cross section of the four censuses and includes period effects. All estimated standard errors are heteroscedasticity robust. Column (1) presents the results for urban to urban flows and excludes flows to/from rural areas and abroad. The parameter estimates are

Table 11.2 Basic gravity model incorporating internal and international flows

	Urban-urban	All flows	Urban-urban	Urban-rural	Urban-world
	(1)	(2)	(3a)	(3b)	(3c)
Distance (km)	−0.811***	−0.810***	−0.811***	n/a	n/a
	(0.0180)	(0.0180)	(0.0180)		
Population (origin)	0.927***	0.937***	0.926***	0.689***	1.613***
	(0.0148)	(0.0150)	(0.0148)	(0.0376)	(0.126)
Population (destination)	0.889***	0.890***	0.889***	0.537***	1.263***
	(0.0153)	(0.0150)	(0.0153)	(0.0445)	(0.0540)
Constant	−9.902***	−10.05***		−9.932***	
	(0.230)	(0.230)		(0.230)	
Year effects	Yes	Yes		Yes	
Obs.	6240	6880		6880	
R^2	0.822	0.861		0.866	
Loglikelihood	−6959.5	−7971.0		−7846.5	

Notes: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Standard errors robust to heteroscedasticity in parentheses

roughly similar to those of Zipf's (1946) original gravity model of the P1 P2/D hypothesis of the intercity movement of persons in the USA, but the hypothesis of coefficients equal to unity is statistically rejected. This extremely simple model fits the data remarkably well: R^2 is 0.822.

Specifications like this are typical in the migration literature, but cross-border and rural-urban flows are excluded in specifications such as Column (1) even though they can be a substantial part of gross migration flows in many countries. Column (2) incorporates these flows but assumes that the origin and destination population effects are the same as in the inter-urban migration specification of Column (1). Given that distance is undefined for flows between urban areas and abroad, or between urban and rural areas, the distance effect is entirely identified from inter-urban flows.[15]

In the regression of Columns (3a)–(3c), we estimate flow-type-specific slopes. This allows us to differentiate the effects of population scale on inter-urban, rural-urban and international-urban flows. Column (3a) shows the population and distance coefficients for inter-urban flows; Column (3b) shows the population coefficients for flows between urban and rural areas; and Column (3c) shows the population coefficients for emigration and immigration between the rest of the world and New Zealand urban areas. As explained previously, the distance coefficient is not defined in Columns (3b) and (3c).

[15]The slight difference in the coefficient of distance in Column (2) as compared with Column (1) is entirely due to the additional 640 (6880–6240) observations (representing migration between urban areas and overseas or rural areas), which are assumed to be affected by population in an identical way as the interurban flows. Once the population impact is allowed to differ between interurban and other flows, the distance effect is again the same as in column (1), as can be seen in column (3).

As expected, given that the coefficients refer exclusively to inter-urban flows in both cases, Column (3a) shows virtually identical results to Column (3c). However, columns (3b) and (3c) shows that population scale effects on urban-rural and urban-world flows are very different from those on inter-urban flows. The much smaller population coefficients for urban to rural and rural to urban flows (0.689 and 0.537, respectively) indicate that rural-urban spatial interaction in the form of migration occurs disproportionally in smaller urban areas. Hence, rural to urban migrants and urban to rural migrants favour relatively smaller urban areas. On the other hand, the much larger coefficients for migration from urban to world and from world to urban areas (1.613 and 1.263, respectively) demonstrate that international migration is disproportionally oriented towards the largest urban areas, i.e. larger areas are preferred either as destinations for immigrants or serve as origins for emigrants.

In Table 11.3, we extend the basic model with additional determinants of migration and again present both pooled and flow-type-specific slopes to identify differences in the impact of these determinants with respect to inter-urban, rural and international flows. Before we discuss the results, it is important to note that, in models of aggregate migration flows that include both locational and population composition characteristics, it is often difficult to disentangle the effect of population composition on mobility from population composition being a push or pull factor for migration. As Plane and Heins (2003) note with respect to differences between age groups in migration systems, it is often better to statistically analyse migration flows of more homogenous sub-groups of the population. However, the macro perspective of estimating determinants of aggregate migration flows still remains informative regarding the main drivers of demographic change in cities and regions, so we will continue to take this approach.

11.4.1 Geographic Determinants

The geographic determinants are distance, a dummy variable measuring interaction between agglomerations and an interisland dummy. These variables can only be linked to inter-urban migration. When comparing their coefficient estimates in which slopes are assumed common across the different types of flows (Column (1)) with the corresponding coefficients in the regression with group-specific slopes (Column (2a)), we see that the coefficients are quite similar. The distance deterrence effect is present in Table 11.3 as before, but the coefficient is less negative than in Table 11.2. Mechanisms that have been suggested in the literature to account for the robustness of the distance deterrence effect include the diminishing information hypothesis (see Schwartz 1973), intervening opportunities hypothesis (see Stouffer 1960; Denslow and Eaton 1984) and the psychic cost hypothesis (see Greenwood 1997; McCann et al. 2010).

The interisland dummy is used to examine what effect, if any, the Cook Strait plays in inter-urban migration flows in New Zealand. The coefficients on the island dummy are negative and statistically significant at the 1% level. It implies that there

Table 11.3 Modified gravity model of internal and international migration flows

		Pooled	Group-specific slopes		
		Common slopes	Urban-urban	Urban-rural	Urban-world
		(1)	(2a)	(2b)	(2c)
Geographic determinants					
Distance (km)		−0.758***	−0.756***	n/a	n/a
		(0.0215)	(0.0216)		
Agglomeration dummy		0.231**	0.265***	n/a	n/a
		(0.0725)	(0.0710)		
Interisland dummy		−0.310***	−0.319***	n/a	n/a
		(0.0377)	(0.0377)		
Demographic determinants					
Population (lagged)	Origin	0.941***	0.923***	0.785***	1.753***
		(0.0214)	(0.0215)	(0.0647)	(0.186)
	Dest.	0.915***	0.932***	0.740***	1.029***
		(0.0204)	(0.0208)	(0.0603)	(0.0527)
Youthfulness (lagged)	Origin	1.256***	1.278***	1.643**	−1.538
		(0.304)	(0.319)	(0.580)	(2.396)
	Dest.	0.0622	−0.420	1.616*	0.0856
		(0.294)	(0.302)	(0.647)	(0.487)
Agedness (lagged)	Origin	−1.951***	−1.720***	1.415	−6.780*
		(0.492)	(0.516)	(0.935)	(3.295)
	Dest.	−1.628**	−1.327*	1.635	0.402
		(0.498)	(0.524)	(0.941)	(1.242)
Prop. of foreign born (lagged)	Origin	−2.170***	−2.129***	−1.461	−5.062*
		(0.267)	(0.269)	(0.787)	(2.513)
	Dest.	−2.655***	−2.919***	−3.473***	2.049***
		(0.265)	(0.267)	(0.909)	(0.590)
International migration effectiveness ratio (lagged)	Origin	−0.00629*	−0.00656*	0.00857	−0.0227
		(0.00256)	(0.00257)	(0.00647)	(0.0279)
	Dest.	0.0145***	0.0132***	0.0158*	0.00731
		(0.00176)	(0.00181)	(0.00619)	(0.00461)
Economic determinants					
Real income growth	Origin	−1.443***	−1.384***	−0.192	−5.591
		(0.305)	(0.309)	(0.809)	(3.446)
	Dest.	1.529***	1.767***	0.297	3.787***
		(0.287)	(0.298)	(0.825)	(0.601)
Inequality growth	Origin	0.646***	0.683**	0.0885	−3.196
		(0.195)	(0.247)	(0.383)	(1.667)
	Dest.	1.830***	0.952***	−0.130	1.026*
		(0.276)	(0.231)	(0.409)	(0.400)

(continued)

Table 11.3 (continued)

		Pooled	Group-specific slopes		
		Common slopes	Urban-urban	Urban-rural	Urban-world
		(1)	(2a)	(2b)	(2c)
Prop. with high educ. (lagged)	Origin	0.503	0.627*	−1.408*	−0.223
		(0.327)	(0.313)	(0.712)	(2.027)
	Dest.	0.818*	1.094***	−1.412	1.708**
		(0.322)	(0.323)	(0.883)	(0.561)
Prop. in own home (lagged)	Origin	−1.173***	−1.142***	−2.093***	−6.425***
		(0.253)	(0.275)	(0.522)	(1.777)
	Dest.	−1.423***	−1.893***	−2.544***	−2.308***
		(0.273)	(0.279)	(0.556)	(0.604)
Climatic determinants					
Rainfall	Origin	−0.113*	−0.138**	0.00119	0.740**
		(0.0531)	(0.0529)	(0.116)	(0.278)
	Dest.	−0.291***	−0.308***	0.0613	−0.113
		(0.0495)	(0.0510)	(0.109)	(0.101)
Sunshine	Origin	−0.737***	−0.681***	0.0386	−4.443**
		(0.173)	(0.176)	(0.399)	(1.468)
	Dest.	0.0262	−0.160	0.240	0.420
		(0.168)	(0.173)	(0.373)	(0.271)
Constant		0.219		2.073	
		(2.054)		(2.145)	
Year effect		Yes	Yes	Yes	Yes
Obs.		6880		6880	
R^2		0.882		0.892	
Loglikelihood		−7402.9		−7114.8	

Notes: $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. Standard errors robust to heteroscedasticity in parentheses

is less migration across the two islands than within them. This result is not surprising given that Cook Strait between the islands presents a big geographical barrier to movement across the islands. Although only about 80 km in length, the journey across the islands can only be carried out via relatively expensive air transportation or a 3-hour journey on a ferry. The agglomeration dummy tests whether there is more migration between the urban areas that make up New Zealand's six biggest cities than between the other urban areas.[16] The coefficient estimate suggests that migration between the six main urban areas is about 26% greater than other migration flows. This reveals strong network effects between the six biggest centres and is

[16]The six biggest cities considered are Auckland, Christchurch, Wellington, Hamilton, Tauranga and Dunedin.

consistent with Poot's (1986) finding of a relatively large exchange of workers between the four main centres over the 1971–1976 period (Auckland, Wellington, Christchurch and Dunedin).

11.4.2 Demographic Determinants

The magnitude of the coefficients on (log) population is similar to those in Table 11.2. However, in the case of flow-type-specific coefficients, the coefficient on destination population is now slightly larger than on origin population. As in Table 11.2, we find that international flows have the largest urban population elasticities of outward and inward migration followed by inter-urban flows, with rural-urban flows exhibiting the smallest coefficients. Hence, as before, rural-urban migration favours smaller urban places, while international migration favours the metropolitan areas.

As expected, age distribution has pronounced effects on migration flows. The youthfulness variable (ratio of population share aged 15–24 over the population share aged 25–54) has a positive and statistically significant effect on outward migration from urban areas, but, interestingly, this is not the case for emigration from urban areas. There is no evidence of migration going specifically towards more youthful urban areas, except for migration from rural to urban areas. A relatively aged population (measured by the ratio of the share of those aged 55 to 64 over those aged 25 to 54) leads to less migration to other urban areas and even less so to the rest of the world (with statistically significant coefficients of −1.720 in Column (2a) and −6.780 in Column (2c), respectively). Emigration is considerably less from urban areas with relatively aged populations.[17] However, a relatively aged population is not a deterrent for urban to rural migration, perhaps signalling migration to lifestyle blocks and/or retirement in non-urban areas.

The presence of immigrants has also notable effects on gross migration flows. The lagged proportion of foreign born has a negative impact on both outward and inward migration of urban areas. The so-called skating rink hypothesis suggests that the foreign born and native born are competing for jobs. A large proportion of foreign born would then suggest greater competition for jobs and be a deterrent for inward migration of the native born. The negative coefficient on the proportion of foreign born in the destination, −2.919 in Column (2a), is consistent with this idea. However, we find that a larger proportion of foreign born leads to less outward migration to other urban areas (with a coefficient of −2.129 in Column (2a)). The negative effect of the foreign born on emigration is even greater (−5.062, see Column (2c)) but is only statistically significant at the 10% level.[18] On the other

[17]This effect is statistically significant at the 1% level in a regression that excludes the observations with zero gross migration.

[18]In a regression that excludes the observations with zero gross migration, the effect is statistically significant at the 1% level, but the coefficient decreases to −2.482.

hand, immigration is going to those urban areas that already have a large proportion of foreign born (a positive coefficient of 2.049 in Column 2(c)), presumably because the existence of networks between earlier immigrants and those still in the home countries reduces migration costs (see Carrington et al. 1996). However, an increase in the proportion of foreign born reduces inflows from rural areas.

The literature shows that evidence of migration responses of the locally born to immigration has been mixed, and these results reinforce that. Frey (1996), Borjas et al. (1996), Borjas (2006) and Glitz (2012) provide evidence that supports the 'skating rink' hypothesis, but Wright et al. (1997), Kritz and Gurak (2001) and Hempstead (2003) found no evidence that an increase in the number of international immigrants lead to the outward migration of the locally born. In fact, Card and DiNardo (2000) argue that increases in the population of immigrants in specific skill groups lead to small increases in the population of native-born individuals of the same skill group. They also conclude that the local labour market impacts of unskilled immigration are mitigated by other avenues of adjustment, such as endogenous shifts in industry structure, rather than through rapid adjustments in the native population.

To further clarify the impact of international migration on inter-urban migration, we also examine the effect of the international migration effectiveness ratio, which is the ratio of the number of immigrants over the number of emigrants in each urban area. It should be noted that this ratio also includes the international movements of the New Zealand born, which are substantial—as noted previously. In New Zealand, immigration exceeds emigration in all urban areas over the 1996–2013 period. Hence the international migration effectiveness ratio is greater than 1 everywhere. Table 11.3 shows that a higher lagged migration effectiveness ratio (i.e. greater net immigration) leads to greater inward inter-urban migration and less outward inter-urban migration, consistent with the urban net migration rate being positively correlated with urban economic growth (see also Ozgen et al. 2010). The effects of the migration effectiveness ratio on rural and international migration are not statistically significant at the 5% level.

11.4.3 Economic Determinants

We examine the respective roles in migration flows of real income growth, changes in the distribution of income, the urban level of human capital and homeownership. As expected, we find that income growth is a significant driver of inter-urban migration flows. Income growth in origin areas leads to less migrant outflows. Income growth in destination areas is associated with more migrant inflows. In order to maximise utility, an individual is expected to offer labour services in the market with the highest expected present value of future wages. Intercensal wage growth is likely to drive wage expectations. Our results are in line with many previous studies, such as, for example, Fan (2005) and Etzo (2008). However, we find that in a developed country like New Zealand, income growth is not a significant factor for migration from urban to rural areas or inflows from rural to urban areas.

Unsurprisingly, we find that income growth is a significant determinant of inflows for international migrants. Immigrants are attracted to areas with high-income growth, although income growth in urban areas does not significantly deter emigration in column 2(c).[19]

Apart from income growth, we examine the role of income inequality on migration (see also, e.g. Stark 2006). The selection model of Borjas (1987) predicts that migration rates are related to inequality in the source and host economies, with, for example, Mayda (2010) and Liebig and Sousa-Poza (2004) providing evidence in support of this at the international level.[20] Phan and Coxhead (2010) provide evidence with internal migration data. Borjas' work implies that an increase in the origin country's relative inequality will have a positive effect on emigration if there is negative selection and a negative effect on emigration if there is positive selection. The relationship between inequality and internal migration is not clear a priori because income inequality can encourage out-migration when a widening of the income gap may have implications for opportunities for people at the bottom of the distribution. These people may be predisposed to migrate to other places with a fairer income distribution. On the other hand, growing inequality can serve as an attractor for migrants who see the growing disparity as an opportunity to make it big in an increasingly winner-take-all environment. Our model includes the growth rate in inequality as measured by the Theil index at the urban area level.[21] We find that inter-urban flows respond to the changes in the distribution of income with a positive relationship between inequality growth and inter-urban flows. An increase in inequality encourages more outflows from, and inflows into, urban areas. This result is likely to reflect the positive correlation between inequality and population size in New Zealand. Previous analysis in New Zealand has found that inequality grew fastest in the largest urban areas (see Alimi et al. 2016), so the positive relationship between inequality growth and migration flows could be a result of larger areas having more migration and higher inequality growth. This could also reflect skill bias in the flows, i.e. positive selection on returns to skills. We find that changes in the income distribution do not significantly affect flows from and to rural areas. For international flows, there is weak evidence that international immigrants are attracted to areas of higher inequality growth (significant at 10% level).[22] This result reflects the draw of the metropolitan areas for immigration, which are also the areas with the greatest inequality growth (Alimi et al. 2016).

[19]Excluding the zero migration flows (corresponding to predominantly smaller urban areas), the coefficient in origin income growth becomes −6.098 and is statistically significant at the 1% level.

[20]Mayda (2010) examined the determinants of migration into 14 OECD countries by country of origin between 1980 and 1995. Liebig and Sousa-Poza (2004) used a dataset that covers over 23 countries, including the countries of the EU, the USA, Canada, Australia and New Zealand.

[21]The Theil index is one of the generalised entropy (GE) measures of inequality. See Conceição and Ferreira (2000) for details.

[22]This effect is no longer statistically significant when observations with zero migration are excluded. However, in the latter case, emigration is lower from urban areas with high inequality, with a coefficient of −1.949 (compare with −3.196 in column 2(c)) that is statistically significant at the 5% level.

The evidence on the impact of human capital (measured by the proportion of the population with at least a bachelor's degree) implies more flows to and from urban areas with a high proportion of people with tertiary education. The 'pull' effect of human capital is larger than the 'push' effect and also statistically stronger (significant at the 1% level rather than the 10% level). Greater outward migration from areas with a high proportion of the highly qualified might simply be capturing the fact that highly qualified people are more migratory (see, e.g. Schwartz 1973; Mincer 1974; McCann et al. 2010). On the other hand, a skilled labour force is an important driver of economic growth in the modern knowledge-driven economy (e.g. Nijkamp and Poot 1998) and thus an attractor of migrants. The large positive coefficient on the human capital variable in immigration to urban areas (1.708, see column (2c)) is not surprising given that New Zealand runs a skill-biased immigration policy. Highly qualified people are likely to settle in places with lots of other highly qualified people to benefit from agglomeration effects, such a matching, sharing and learning (Duranton and Puga 2004). Excluding observations with zero migration (likely to correspond to the smaller urban areas), it can be shown that emigration rates are significantly higher (at the 5% level) from urban areas with a larger fraction of graduates.

The role of homeownership is the final economic variable we examine. We find that home ownership has a negative relationship with both migration inflows and outflows. This result is consistent across inter-urban, urban-rural or urban-international migration flows. The lower mobility in areas with high homeownership could be associated with the increased cost of migration that comes with homeownership, such as the transaction cost of selling the house, the loss of local social capital and the psychic cost of leaving (e.g. Roskruge et al. 2013). A high ownership rate in the destination can also signal a tight rental market to would-be migrants and might discourage migration inflows of renters.

11.4.4 Climatic Determinants

Graves' (1979) study in the USA was one of the earliest studies to find an important role for climate as a determinant of internal migration. Using data from Standard Metropolitan Statistical Areas, Graves examined the effects of temperature, relative humidity and wind velocity on net migration. The study found weather variables to be statistically significant determinants of migration. Without controlling for amenity variables, income is insignificant in determining migration in Graves' research, but income becomes significant and exhibits a life-cycle pattern after controlling for amenity variables. Other studies such as Rappaport (2007) and Poston et al. (2009) have reconfirmed these results for the USA. Evidence from Europe, however, tends to suggest a smaller climate effect on migration (Cheshire and Magrini 2006). However, studies like Hunt (1993) and Evans (1993) argued for a limited role of climate in determining migration.

Poot (1986) examined the effects of average temperature and rainfall in a New Zealand urban area and found that warmer urban areas had lower out-migration. Additionally, migrants were relatively more attracted to the wetter urban areas of the North Island. We provide here new evidence on the effect of climate variables in the New Zealand context. As noted in the previous section, the climate variables available for this research are the 30-year average of sunshine hours in each urban area and the 30-year average of rainfall measured in millimetres. Here, we find less inter-urban outflows and inflows to/from wet areas but higher international emigration flows from wet urban areas (perhaps an Auckland effect: Auckland has relatively high rainfall). However, rainfall is not a significant determinant for rural to urban migration. Sunshine is not an important pull factor (perhaps due to only considering migration of those aged 25–54 and not retirement migration) but appears to deter inter-urban outflows. While rainfall and sunshine are statistically significant predictors of urban emigration in Column 3(c) (with coefficients of 0.740 and −4.443, respectively), it can be shown that removing the observations with zero gross migration makes the coefficients statistically insignificant.

11.5 Conclusion

In this study, we revisited the modelling of gross internal migration in New Zealand. Using pooled data on gross migration between 40 urban areas in 4 successive intercensal periods, we identified a range of geographic, demographic, economic and climatic characteristics of urban areas that are statistically significant determinants of migration. New Zealand is a small but open population in which one quarter of the resident population is foreign born and one sixth of the New Zealand population lives abroad. In this context, we show that it is important to model inter-urban migration jointly with rural-urban migration and with international migration. We find notable differences in the roles of migration determinants when comparing urban-urban, urban-rural and urban-world migration flows. The estimation of these models is straightforward and can be easily applied to other case studies in which international and/or rural-urban migration are the important components of population churn.

Not surprisingly, migration in New Zealand responds to factors that have shown to be highly robust determinants of migration in the literature. We find evidence that migrants are attracted to areas with growing income. Positive growth in income deters migration outflows. However, we also find a positive relationship between inequality growth and migration flows.

With respect to demographic characteristics, population age composition is very important for outward migration. As expected, youthfulness increases outward migration while agedness decreases it. The effect of age composition of destination areas is less conclusive. We find no evidence to support the skating rink hypothesis that immigration triggers outward internal migration. In fact, the opposite is the case: internal migrants move to urban areas where international migrants settle. With respect to the role of climate, we find that sunnier areas have less out-migration, but the effect of rainfall is mixed.

The use of models of aggregate gross migration flows, such as modified gravity models, presents also some challenges, especially when interpreting the coefficients. There is difficulty in distinguishing whether the results from the models are related to average personal characteristics of individuals in the area, with heterogeneity in mobility linked to such characteristics or whether average personal characteristics of an area are push factors or pull factors for migration. This problem has been well-established in the literature and has triggered the increasing use of microdata in migration research in recent decades. However, aggregate models and data like those employed in this chapter still have a role to play in regional science, especially when the focus is on understanding subnational socio-economic trends.

There are several ways in which the research here can be extended. First, the gross migration model can be embedded in a dynamic spatial equilibrium model in which the endogeneity of several of the migration determinants used in the present study is explicitly taken into account.[23] Second, there is considerable heterogeneity and selection in the population with respect to migration behaviour, and this can impact on estimated responses to migration incentives. Third, the model in the present chapter assumes temporal stability of regression coefficients in models of gross migration but that is rarely tested (and testing it with our data is beyond the scope of the chapter). Fourth, where data are available on specific origins and destinations of international migrants, these flows can be modelled jointly with the internal migration flows. To date, this has only been done in the New Zealand context with respect to trans-Tasman migration between Australia and New Zealand (see Poot 1995; Gorbey et al. 1999). Furthermore, it is clear that spatial spillovers and spatial heterogeneity have not been taken into account in the estimates presented in this study even though these can be potentially important (see, e.g. LeSage and Pace 2008, 2009; Peeters 2012). Finally, we may expect that migration research may start to exploit new sources of data such as integrated administrative data or data from social networks.

Hence, research on internal migration is by no means a dead-end area of social science. We can be confident that further advances will be made in the coming years in modelling migration and other forms of spatial interaction. Such new developments are likely to take account of spatial sorting, heterogeneity and spatial spill-overs in a general equilibrium setting. In such models, research can account for any non-trivial form of spatial frictions and may also exploit the rich new microdata that sources of 'big data' are now starting to provide.

[23]See, e.g. for the USA, Blanchard and Katz (1992); Europe, Decressin and Fatás (1995); Australia, Debelle and Vickery (1999); and New Zealand, Choy et al. (2002).

Acknowledgements Research for this study was originally funded by the 2012–2014 *Nga Tangata Oho Mairangi* project, Ministry of Business Innovation and Employment grant CONT-29661-HASTR-MAU, and more recently by the 2016–2019 *National Science Challenge 11: Building Better Homes, Towns and Cities* project. Access to the data used in this study was provided by Statistics New Zealand (SNZ) under conditions designed to give effect to the security and confidentiality provisions of the Statistics Act 1975. All frequency counts using census data were subject to base three rounding in accordance with SNZ's release policy for census data.

References

Alimi O, Maré DC, Poot J (2016) Income inequality in New Zealand regions. In: Spoonley P (ed) Rebooting the regions: why low or zero growth needn't mean the end of prosperity. Massey University Press, Auckland, pp 177–212

Blanchard O, Katz LF (1992) Regional evolutions. Brook Pap Econ Act 1:1–75

Borjas GJ (1987) Self-selection and the earnings of immigrants. Am Econ Rev 77(4):531–553

Borjas GJ (2006) Native internal migration and the labor market impact of immigration. J Hum Resour 41(2):221–258

Borjas GJ, Freeman RB, Katz LF (1996) Searching for the effect of immigration on the labor market. Working Paper Number w5454. National Bureau of Economic Research

Caldwell JC (1968) Determinants of rural-urban migration in Ghana. Popul Stud 22(3):361–377

Cameron MP, Poot J (2019) The estimation and interpretation of coefficients in panel gravity models of migration. Lett Spat Resour Sci 12(1):9–15

Card D, DiNardo JE (2000) Do immigrant inflows lead to native outflows? Working Paper Number w7578). National Bureau of Economic Research

Carrington WJ, Detragiache E, Vishwanath T (1996) Migration with endogenous moving costs. Am Econ Rev 86(4):909–930

Cebula RJ, Alexander GM (2006) Determinants of net interstate migration, 2000–2004. J Reg Anal Policy 36(2):116–123

Cheshire PC, Magrini S (2006) Population growth in European cities: weather matters-but only nationally. Reg Stud 40(1):23–37

Choy WK, Maré DC, Mawson P (2002) Modelling regional labour market adjustment in New Zealand. Working Paper No. 02/01. New Zealand Treasury

Conceição P, Ferreira P (2000) The young person's guide to the Theil index: suggesting intuitive interpretations and exploring analytical applications. UTIP Working Paper No. 14. University of Texas

Curran SR, Rivero-Fuentes E (2003) Engendering migrant networks: the case of Mexican migration. Demography 40(2):289–307

Debelle G, Vickery J (1999) Labour market adjustment: Evidence on interstate labour mobility. Aust Econ Rev 32(3):249–263

Decressin J, Fatás A (1995) Regional labor market dynamics in Europe. Eur Econ Rev 39 (9):1627–1655

Denslow D, Eaton PJ (1984) Migration and intervening opportunities. South Econ J 51(2):369–387

Duranton G, Puga D (2004) Micro-foundations of urban agglomeration economies. In: Henderson JV, Thisse J-F (eds) Handbook of regional and urban economics, vol 4. Chapter 48, pp 2063–2117

Esipova N, Pugliese A, Ray J (2013) The demographics of global internal migration. Migrat Policy Pract J 3(2):3–5

Etzo I (2008) Determinants of interregional migration in Italy: a panel data analysis. J Reg Sci 51 (5):948–966

Evans AW (1993) Interregional equilibrium: a transatlantic view. J Reg Sci 33(1):89–97

Fan CC (2005) Modeling interprovincial migration in China, 1985–2000. Eurasian Geogr Econ 46 (3):165–184

Frey WH (1996) Immigration, domestic migration and demographic Balkanization in America: new evidence for the 1990s. Popul Dev Rev 22(4):741–763
Glaeser E (2012) The triumph of the city: how our greatest invention makes us richer, smarter, greener, healthier, and happier. Penguin Press, New York
Glitz A (2012) The labor market impact of immigration: a quasi-experiment exploiting immigrant location rules in Germany. J Labor Econ 30(1):175–213
Gorbey S, James D, Poot J (1999) Population forecasting with endogenous migration: an application to trans-Tasman migration. Int Reg Sci Rev 22(1):69–101
Graves PE (1979) A life-cycle empirical analysis of migration and climate, by race. J Urban Econ 6 (2):135–147
Greenwood MJ (1969) An analysis of the determinants of geographic labor mobility in the United States. Rev Econ Stat 51(2):189–194
Greenwood MJ (1975) Research on internal migration in the United States: a survey. J Econ Lit 13 (2):397–433
Greenwood MJ (1997) Internal migration in developed countries. In: Rosenzweig MR, Stark O (eds) Handbook of population and family economics. Volume 1B. Elsevier, Amsterdam, pp 647–720
Hempstead K (2003) Immigration and native migration in New York City, 1985–1990. Popul Res Policy Rev 22(4):333–349
Hunt GL (1993) Equilibrium and disequilibrium in migration modelling. Reg Stud 27(4):341–349
Hunt GL, Mueller RE (2004) North American migration: returns to skill, border effects, and mobility costs. Rev Econ Stat 86(4):988–1007
Iammarino S, Rodriguez-Pose A, Storper M (2019) Regional inequality in Europe: evidence, theory and policy implications. J Econ Geogr 19(2):273–298
Jeanty P, Partridge M, Irwin E (2010) Estimation of a spatial simultaneous equation model of population migration and housing price dynamics. Reg Sci Urban Econ 40(5):343–352
Karemera D, Oguledo VI, Davis B (2000) A gravity model analysis of international migration to North America. Appl Econ 32(13):1745–1755
King R, Skeldon R (2010) 'Mind the gap!' Integrating approaches to internal and international migration. J Ethn Migr Stud 36(10):1619–1646
Kritz MM, Gurak DT (2001) The impact of immigration on the internal migration of natives and immigrants. Demography 38(1):133–145
LeSage J, Pace RK (2008) Spatial econometric modelling of origin-destination flows. J Reg Sci 86 (3):393–421
LeSage J, Pace RK (2009) Chapter 9: Spatial econometric interaction models. In: Introduction to spatial econometrics. CRC Press, Boca Raton
Liebig T, Sousa-Poza A (2004) Migration, self-selection and income inequality: an international analysis. Int Rev Soc Sci 57(1):125–146
Long LH (1991) Residential mobility differences among developed countries. Int Reg Sci Rev 14 (2):133–147
Maré DC, Choy WK (2001) Regional labour market adjustment and the movements of people: a review. Working Paper No. 01/08. New Zealand Treasury
Maré DC, Timmins J (2000) Internal migration and regional adjustment: some preliminary issues. In: Morrison P (ed) Labour, employment and work (LEW) in New Zealand: proceedings of the LEW9 conference. Geography Department and Industrial Relations Centre, Victoria University of Wellington, Wellington
Maré DC, Morten M, Stillman S (2007) Settlement patterns and the geographic mobility of recent migrants to New Zealand. N Z Econ Pap 41(2):163–195
Maxwell NL (1988) Economic returns to migration: marital status and gender differences. Soc Sci Q 69(1):108–121
Mayda AM (2010) International migration: a panel data analysis of the determinants of bilateral flows. J Popul Econ 23(4):1249–1274
McCann P, Poot J, Sanderson L (2010) Migration, relationship capital and international travel: theory and evidence. J Econ Geogr 10(3):361–387

Michaelides M (2011) The effect of local ties, wages, and housing costs on migration decisions. J Socio-Econ 40(2):132–140
Mincer J (1974) Schooling, experience and earnings. Columbia University Press, New York
Modestino AS, Dennett J (2013) Are American homeowners locked into their houses? The impact of housing market conditions on state-to-state migration. Reg Sci Urban Econ 43(2):322–337
Nijkamp P, Poot J (1998) Spatial perspectives on new theories of economic growth. Ann Reg Sci 32 (1):7–37
Ozgen C, Nijkamp P, Poot J (2010) The effect of migration on income growth and convergence: meta-analytic evidence. Pap Reg Sci 89(3):537–561
Pedersen PJ, Pytlikova M, Smith N (2008) Selection and network effects: migration flows into OECD countries 1990–2000. Eur Econ Rev 52(7):1160–1186
Peeters L (2012) Gravity and spatial structure: the case of interstate migration in Mexico. J Reg Sci 52(5):819–856
Phan D, Coxhead I (2010) Inter-provincial migration and inequality during Vietnam's transition. J Dev Econ 91(1):100–112
Pissarides CA, McMaster I (1990) Regional migration, wages and unemployment: Empirical evidence and implications for policy. Oxf Econ Pap 42(4):812–831
Plane DA (1993) Demographic influences on migration. Reg Stud 27(4):375–383
Plane DA, Heins F (2003) Age articulation of U.S. inter-metropolitan migration flows. Ann Reg Sci 37(1):107–130
Poot J (1986) A system approach to modelling the inter-urban exchange of workers in New Zealand. Scott J Polit Econ 33:249–274
Poot J (1995) Do borders matter? A model of interregional migration in Australasia. Australas J Reg Stud 1(2):159–182
Poot J, Alimi O, Cameron M, Maré D (2016) The gravity model of migration: the successful comeback of an ageing superstar in regional science. Invest Reg J Reg Res 36:63–86
Poston DL, Zhang L, Gotcher DJ, Gu Y (2009) The effect of climate on migration: United States, 1995–2000. Soc Sci Res 38(3):743–753
Ramos R (2016) Gravity models: a tool for migration analysis. IZA World Labor. https://doi.org/10.15185/izawol.239
Rappaport J (2007) Moving to nice weather. Reg Sci Urban Econ 37(3):375–398
Ravenstein EG (1885) The laws of migration, part 1. J R Stat Soc 48(2):167–235
Ravenstein EG (1889) The laws of migration, part 2. J R Stat Soc 52(2):241–305
Roskruge M, Grimes A, McCann P, Poot J (2013) Homeownership, social capital and satisfaction with local government. Urban Stud 50(12):2517–2534
Schwartz A (1973) Interpreting the effect of distance on migration. J Polit Econ 81(5):1153–1169
Sjaastad LA (1962) The costs and returns of human migration. J Polit Econ 70(5):80–93
Sloan M (2013) The post-move satisfaction of individuals moving within New Zealand. Doctoral thesis, Victoria University of Wellington, Wellington
Stark O (2006) Inequality and migration: a behavioral link. Econ Lett 91(1):146–152
Statistics New Zealand (2013). *2013 Census QuickStats about national highlights.* Retrieved from www.stats.govt.nz
Stouffer SA (1960) Intervening opportunities: a theory relating mobility and distance. J Reg Sci 2 (1):1–26
Westerlund O (1998) Internal migration in Sweden: the effects of mobility grants and regional labour market conditions. Labour 12(2):363–388
Wright RA, Ellis M, Reibel M (1997) The linkage between immigration and internal migration in large metropolitan areas in the United States. Econ Geogr 73(2):234–254
Zelinsky W (1971) The hypothesis of the mobility transition. Geogr Rev 61(2):219–249
Zipf GK (1946) The P1 P2/D hypothesis: on the intercity movement of persons. Am Sociol Rev 11 (6):677–686

Chapter 12
Baby Boomers' Paths into Retirement

Ayoung Kim and Brigitte S. Waldorf

Abstract Baby boomers—comprised of persons born between 1946 and 1964—made up almost one-third of the US population in 2010, the year that the first baby boomers reached retirement age. The baby boom generation is increasingly dominating the older segment of the population. However, we have not yet thoroughly analyzed whether the baby boom generation will follow the migration preferences of older cohorts. This research focuses on how the nexus of aging, migration, and income plays out for the older population, with a particular emphasis on the baby boom generation. The results show a distinct North-South disparity through spatial income redistribution due to elderly migration. We find that—in terms of net income gains—the winners of retirement migration are states located in the West and South, whereas the interior and the East Coast are losing net income. Older persons' migration across interstate boundaries creates this stark income gain disparity because it is a highly selective process. Those choosing to move out of state are predominantly white, affluent, and better educated persons. Importantly, we also find a strong positive association between income and interstate moves for the baby boomers, but not for the WWII generation. The income selectivity and the associated spatial income redistribution will become stronger in the future when baby boomers become more and more dominant among the older population

Keywords Elderly migration · Baby boomers · Migration determinants · Spatial income redistribution

A. Kim
Department of Agricultural Economics, Mississippi State University,
Mississippi State, MS, USA
e-mail: ayoung.kim@msstate.edu

B. S. Waldorf (✉)
Department of Agricultural Economics, Purdue University, West Lafayette, IN, USA
e-mail: bwaldorf@purdue.edu

R. S. Franklin (ed.), *Population, Place, and Spatial Interaction*, New Frontiers in Regional Science: Asian Perspectives 40,
https://doi.org/10.1007/978-981-13-9231-3_12

12.1 Introduction

The US population is getting older. Rising life expectancies in combination with low fertility are leading to a rapidly growing group of older Americans, both in absolute and relative terms (Waldorf and McKendree 2013). Compared to other developed countries, the USA takes on a unique position as the *share* of the older generation is relatively small—only 16% compared to about 22% in Germany and Italy and more than 28% in Japan in 2018.[1] However, the absolute growth of the older generation in the USA is huge and unprecedented because the baby boom generation is exceptionally large. Baby boomers—comprised of persons born between 1946 and 1964—made up almost one-third of the US population in 2010, the year when the first baby boomers reached retirement age (65 years old). By now, the oldest baby boomers are in their 70s and many already retired, refocusing their lives away from work. For some, this adjustment involves moving and downsizing, with locational preferences often centered about amenities. The implication is that—as many pointed out years ago—we see a rising number of older migrants (Longino and Bradley 2003; Plane and Jurjevich 2009).

From a place-based perspective, local governments and community developers are interested in senior migration as an economic development strategy for quite some time (Serow 2003). In receiving communities, seniors purchase local goods and services using externally derived income from Social Security, private pensions, and certain forms of equity income. In addition, the elderly drive the demand for additional government services or enhance a local government's economy and tax base without imposing heavy demands on local services. Particularly, in regions and communities that are not able to attract big manufacturing firms, this migration-induced employment effect is seen as a way to boost community businesses and create vibrant economic places. The other side of the coin is, of course, that communities that lose their elderly population may also lose business opportunities. Moreover, as Davies (2014) argues, in preferred destination communities, the inflow of older people can also induce disadvantages such as rising housing prices and congestion. And these disadvantages can be further exaggerated when taking into account seasonal elderly migration (Smith and House 2006).

How much communities can gain—or lose—due to senior migration depends on migrants' incomes and wealth. In the aggregate, we see substantial variations in income gains and losses resulting from elderly migration. For example,[2] New Hampshire, North Dakota, and Wisconsin not only lost a substantial portion of their elderly population due to migration but also the net migration losses among well-off seniors were disproportionally high. Florida, on the other hand, had an overall positive net migration among the elderly; moreover, the net migration rate for seniors with an annual income exceeding US$70,000 is higher than for any other income group.

---

[1] US Census Bureau, International Data Base, https://www.census.gov/data-tools/demo/idb/informationGateway.php, accessed March 14, 2019.

[2] Based on the IPUMS-USA data: ACS 2006–2010 (Ruggles et al. 2018)

In this chapter, we focus on the income dependency of seniors' migration behaviors and migration patterns, combining a microlevel analysis of who moves and where people move, with a macrolevel analysis of the aggregate outcomes. At the microlevel, we estimate how income affects the migration propensity of the elderly, as well as how it affects their destination choices. Moreover, we ask whether the income dependency of senior migration is modulated by cohort and gender effects. These are important questions given salient demographic shifts and societal changes. First, the baby boom generation will increasingly dominate the older segment of the population, and we have not yet fully analyzed (or even fully observed) whether the baby boom generation will follow the migration preferences of older cohorts (Haas and Serow 2002). Second, women's life expectancy exceeds that of men at all ages, leading to sex imbalances with women making up the majority of the older population. Furthermore, over time, women have gained some economic power and have become more likely to enter retirement with higher incomes and wealth. At the macrolevel, we investigate the aggregate outcomes of senior migration, utilizing migration flow data by age groups. Most importantly, we translate the senior migration flows into flows of income so as to identify regions where senior migration may serve as an economic development strategy.

The remainder of this chapter is organized as follows: after this introduction, the next section provides some background on the aging US society and previous research on later-life migration. This is followed by the empirical section, which includes the research design, methods, data, and results. The chapter ends with a summary and conclusion.

12.2 Background

12.2.1 Aging Society

Since the turn of the century, the US population age 65 and older grew by almost 20 million people or 57% from 32.6 million in 2000 to 51.1 million in 2018 (Fig. 12.1). This growth in absolute terms is accompanied by substantial relative growth. In 2000, the 65+ population accounted for merely 11.9% of the total population but increased to 15.8% in 2018. These substantial changes in the age composition of the US population are the beginning of what is expected to become an unprecedented aging boom. The US Census Bureau predicts that by 2030—the year when all baby boomers have entered retirement age—20% of the US population will be 65 or older. By 2035, persons aged 65+ will outnumber the young population under the age of 18 (U.S. Census Bureau 2018).

While we have a good understanding of the causes of population aging—in short, population aging is the outcome of past fertility decline and fluctuations, coupled with ongoing mortality decline, and modulated by the influx of mostly young immigrants—the consequences are not yet fully understood. Clearly, the growing number of older residents will increase the demand for elder-oriented goods and

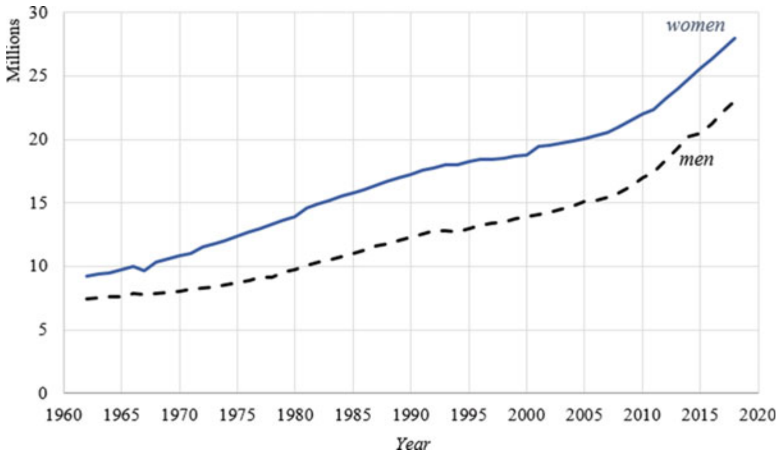


Fig. 12.1 US population aged 65 and above, 1961–2018. (Source: US CPS data, accessed via IPUMS-CPS)

services such as doctors, hospitals, transportation services, and housing that is accessible despite mobility limitation. Yet, there are a number of developments that make it difficult to predict the magnitude of the increasing demand and the magnitude of a possible bottleneck in the provision of such services. One such development is that people choose to retire at an older age. There are two main reasons for such a retirement postponement (Kromer and Howard 2013; Sewdas et al. 2017): (1) many people stay in good health well beyond the age of 65 years; and (2) financial constraints force some people to retire later in life. Another development is the feminization of the older population (Waldorf and Pitfield 2005) and the changing role of women. Women, on average, outlive men and make up the majority of the older population. Furthermore, the gender wage gap is slowly declining, women are increasingly likely to participate in the labor force and get better access to traditionally male-dominated occupations and positions. In the years ahead, therefore, women will enter retirement with more economic power than earlier cohorts. As such, they may also be less able and willing to fulfill the caretaker role for aging parents—a role that traditionally was assigned almost exclusively to women.

Even if good nationwide estimates of the demand for elder-oriented goods and services are available, regional estimates are difficult to obtain, given that many people change location upon retirement. From a regional economic perspective, the redistribution of the older population can cause demand or supply shortages of elder-oriented services. Moreover, whether the postretirement redistributions are advantageous for a region depends on the economic power of retirees received or retained. The winners will be those regions that can attract and retain the affluent retirees who spend their wealth and income in local communities and do not demand many services from the communities.

12.2.2 What Do We Already Know About Later-Life Migration?

Four decades ago, Wiseman and Roseman (1979) already drew attention to much-needed research on elderly migration, arguing that an elderly population is a heterogeneous group and that elderly migration behavior is not adequately captured in general migration models. They developed a typology of elderly migration. Types of moves that are of particular relevance to older people are, for example, kinship moves and returns to the place of birth. Although the typology focuses on the microscale of individual migration behavior, Wiseman and Roseman (1979) also recognized that elderly migration has the power to distinctly alter communities, and thus takes on a pivotal role for planners and policymakers.

Ever since Wiseman and Roseman's seminal contribution, the literature on senior migration has grown substantially, and we have gained a good understanding of later-life mobility behavior and patterns. Walters (2002) provides an excellent review of studies in the 1990s. More recent studies, by and large, confirm the earlier trends. In particular—as Clark et al. (1996) already concluded—both personal and locational characteristics are important factors determining elderly migrants' decisions to change their state of residence. A recurring theme is the age dependency of older person's migration propensities and locational choices, in particular, the distinction between younger and older elderly.

Taking a life-time perspective on migration, Plane and Heins (2003) show that the decision to move as well as the destination choice systematically vary according to a person's stage in the life cycle. A persistent empirical regularity is a slight hump in migration propensity as people reach retirement age, a feature that Rogers (1988) had repeatedly alluded to in his migration age schedules. For elderly migrants, Plane and Heins (2003) further find a distinction between retirement migrants heading to traditional retirement areas like Florida and Arizona, and older seniors who are originating in the sunbelt and returning, presumably to be with family. Plane and Jurjevich (2009) show that the redistribution of the older population goes beyond a spatial concentration from many origins into fewer destinations, but also includes a hierarchical component. That is, the younger elderly move down the urban hierarchy from large metro areas to amenity-rich smaller metro or even rural areas. In contrast, the older elderly may very well follow their adult children and/or make healthcare motivated moves up the urban hierarchy.

Similar linkages between senior migration and locational/personal characteristics—especially age—are repeatedly found in the literature. Conway and Houtenville (2001, 2003) emphasize the importance of taxes for seniors' destination choices. They find evidence for a difference in migration behavior between younger and older seniors, and relate it to differential responses to taxes and cost of living. Somenahalli and Shipton (2013) find that the locational choices of people in their mid-70s are associated with an area's access to services, whereas socio-economic indicators are more relevant for younger and more recent retirees.

Among location attributes, unquestionably, the most attention is directed toward the role of amenities. The common view is that—as people retire from the labor force and become footloose—they follow their locational preferences and move to amenity-rich areas (Duncombe et al. 2001). Whisler et al. (2008) find that older college-educated persons show a strong preference for mild climates. However, realizing such locational preferences requires what Walters (2000) refers to as "enabling attributes," such as having sufficient income to take advantage of high-end cultural offerings. Natural amenities—a warm climate, for example—might seem more accessible for everyone, but if amenity compensation (Graves and Linneman 1979) occurs primarily through high housing prices rather than low wages, then low-income retirees may find themselves excluded.

Income is also of pivotal importance for the macrolevel consequences of elderly migration. Day and Barlett (2000) document the positive economic impact of wealthy retirees in struggling counties in Texas. However, Serow (2001, 2003) concludes that there is some sorting whereby the well-off retirees sort into prosperous locations. Moreover, the elderly population in retirement communities changes over time, not only with respect to their health status and their needs but also their financial means and spending patterns. In light of these changes, Rowles and Watkins (1993) suggested that attracting elderly migrants as a development strategy needs to consider the long-term implications. Serow's (2003) review paper documents that elderly in-migrants are indeed a boost to the local economy in the short run. However, he also points to the need for long-term analyses.

12.2.3 How Do We Expect Later-Life Migration to Change?

We expect later-life migration to change in light of deep transformations affecting the population at risk of contemplating to migrate at a later age. The US population of age 65 and above is growing both in absolute and relative terms, and it will be increasingly made up of baby boomers who have a vastly different life-time experience than earlier generations, that is the cohorts born during WWII and the Great Depression. Moreover, the generation now entering retirement age is expected to live longer than earlier generations, and stay healthy longer.[3] Finally, the share of women entering retirement with economic independence and power will rise.

The empirical analysis thus seeks to answer four questions: (1) What are the aggregate outcomes of later-life migration, especially with respect to income variation and redistribution? (2) How does income affect the migration behavior of the older population? (3) Does the migration behavior's income dependency differ between the baby boom generation and the pre-baby boom generation? (4) Does the migration behavior's income dependency differ between men and women?

[3]Increasing mortality rates in the US as documented by Case and Deaton (2015) are, as of now, mostly confined to whites of age 45–54 years.

12.3 Empirical Analysis

We address the research questions by focusing on testing three main hypotheses for people aged 65 and above in the USA. First, we hypothesize that—compared to lower-income seniors—affluent seniors are more likely to migrate, more likely to move out-of-state, and more likely to move to a traditional retirement destination (TRD). Second, we test whether the income dependency of migration behavior differs between baby boomers and pre-baby boomers. Finally, we test whether there is a gender difference in the income dependency on migration behavior. We use American Community Surveys (ACS) data to estimate a battery of logit models addressing these hypotheses. The empirical analysis includes the translation of microlevel results into aggregate migration patterns and spatial variations.

12.3.1 Model

We estimate logit models for three types of binary migration variables. The first migration variable, M_1, distinguishes between the basic choice between moving and staying. For the subset of movers, the second migration variable, M_2, distinguishes between moving to a different state versus moving within state boundaries. For the interstate movers, we define a third migration variable, M_3, distinguishing between those moving to a traditional retirement destination and those moving elsewhere. For each migration variable, M_c (c = 1,2,3), we estimate several model specifications. The first is a basic model of the form:

$$\text{Prob}(M_c = 1) = \text{logit}(X\beta) \tag{12.1}$$

where X is a matrix of covariates and β is a vector of parameters. The set of **X** variables will include income to test for the hypothesized income dependency. In the second specification, the basic model is reestimated with fixed effects for the (premigration) place of residence. This is an important extension because places may differ by salient variables that affect migration but are difficult to measure, including stability, feelings of belonging, attachment, and cultural norms. Finally, we will allow the parameters to vary by cohort so as to answer the question whether income indeed plays different roles for baby boomers than for earlier generations. The third specification includes interaction terms that yield gender-specific parameters for income dependency.

Table 12.1 US elderly household heads[a] by household structure

Household heads *without* spouse		Household heads *with* spouse (married-couple households)	
190,797,243	Female: 71%	147,506,732	Female: 22%
(56.4%)		(43.6%)	

Source: American Community Survey, 2005–2017, IPUMS-USA
[a]Household heads age 65+; spouses are not age-restricted (78% are 65+)

12.3.2 Data Selection and Data Description

We use data from the annual American Community Surveys (ACS) during the 13-year period, 2005–2017. The ACS is a household survey administered by the US Census Bureau. We accessed the data from Minnesota University's Integrated Public Use Microdata Series (IPUMS-USA), which integrates the annual ACS data into an overall dataset (Ruggles et al. 2018). The ACS data are particularly well suited for the analysis because the questionnaires barely changed during our study period. This allows us to conduct a multi-year analysis that includes different cohorts of elderly persons.

We select all household heads of age 65 or older at the time of the respective surveys. This yields a total number of $n = 4{,}172{,}204$ observations in the sample, representing 338 million household heads over the 13-year period, or 26 million per year on average. As shown in Table 12.1, the majority of household heads (56%) are living without a spouse and thus are widowed, divorced, separated, or never married. Married-couple households are a minority among the older population. Remarkable, albeit not unexpected given men's comparatively high mortality rates, is the gender distribution among these two types of households. Most household heads without a spouse are women (71%). In married-couple households, women are rarely (22%) identified as the household head as can be expected from the more traditional gender roles of older generations.

Table 12.2 shows the generational transformation of the older population. Over time, the share of baby boomers is increasing substantially. In 2011, the oldest baby boomers surpassed the age-65 sample selection threshold for household heads for the first time. In that year, they represented about 5.4% of all households without a spouse. Baby boomer shares kept increasing, and by 2017, they already represented more than 39% of all elderly households without a spouse. Because our sample selection did not impose age restrictions on the spouses, baby boomers can "enter" indirectly in elderly households in which a spouse is present. They are not included as an independent observation, but they undoubtedly are involved in the migration decision-making, either as a younger spouse of somebody from a pre-baby boom generation or as an older spouse of a baby boomer not yet exceeding the 65-year age threshold. Not surprisingly, the share of households where both spouses belong to the baby boom generation is increasing. It is 0% before 2011, but it has already reached 43% just 6 years later. It will take until the 2030s before the baby boomer

Table 12.2 Representation of baby boomers among US household heads (HH) of age 65+

Year	HH *without* spouse		HH *with* spouse		
	Total	Boomers (%)	Total	HH and spouse are boomers(%)	Either HH or spouse is boomer(%)
2005	3,114,994	0	9,887,080	0	8
2006	12,944,981	0	9,793,396	0	10
2007	13,052,008	0	9,979,744	0	11
2008	13,498,895	0	10,154,770	0	15
2009	13,645,117	0	10,507,769	0	18
2010	14,107,650	0	10,804,827	0	22
2011	14,322,538	5	11,101,084	6	23
2012	14,870,071	12	11,622,036	13	23
2013	15,329,961	17	11,887,967	20	23
2014	15,835,596	23	12,287,672	26	23
2015	16,241,916	28	12,763,500	32	22
2016	16,751,278	33	13,116,919	38	21
2017	17,082,238	39	13,599,968	43	21

Source: American Community Survey, 2005–2017, IPUMS-USA
Note: Samples include household heads age 65+; spouses are not age-restricted

share will gradually decrease and be replaced by the post-baby boom cohort, that is, by household heads born after 1964.

Table 12.3 lists the summary statistics of the relevant variables in our analysis. The upper part includes the variables of interest, that is, the three migration variables, M_1, M_2, and M_3 (see Sect. 12.3.1). M_1 is defined for all household heads, and it takes on the value 1 for those who moved during the last year, referred to as "mover." Everybody else is referred to as "stayer." It is important to stress, however, that the stayers may have moved at some earlier point in time. Overall, 5.1% of all household heads are defined as movers. The variable M_2 is defined for movers only and takes on the value 1 if the household moved into a different state (interstate move). In total, only 18.6% of senior moves are interstate moves. Lastly, the variable M_3 is defined for interstate movers and takes on the value 1 if the destination state is a traditional retirement destination, defined as Arizona, Florida, Georgia, North Carolina, South Carolina, and Texas. Almost 40% of interstate movers are headed to a traditional retirement destination.

The lower part of Table 12.3 shows the characteristics of household heads age 65 or older, both overall and separately for movers ($M_1 = 1$) and stayers ($M_1 = 0$). The baby boomers make up 15.8% of the elderly household heads and are slightly overrepresented among the movers. On average, movers and stayers are equally old. The median age, however, is 1 year younger for movers than for stayers. Interestingly, compared to their spouses (if present), household heads are about 4 years older both for moving and staying households.

Female household heads account for 49.8% overall, but they are overrepresented among movers where women account for 53.5%. Many of the female movers may very well be women who outlived their husbands and moved shortly after becoming

Table 12.3 Summary statistics

Dependent variable		Mean		
M_1 (base: all HH age 65+)		0.051		
M_2 (base: all HH age 65+ with $M_1 = 1$)		0.186		
M_3 (base: all HH age 65+ with $M_2 = 1$)		0.391		
Characteristics of household heads age 65+		**Total**	**Stayer**	**Mover**
Baby boomers (%) (born between 1946 and 1964)		15.8	15.7	17.5
Female (%)		49.8	49.6	53.5
Mean age (years)		74.9	74.9	74.9
Family structure				
Married (with spouse) (%)		43.6	44.2	32.7
Spouse's mean age (years)		70.4	70.4	69.8
Living without a child (%)		87.1	86.9	90.9
Living with a child aged 19 or older (%)		12.4	12.6	8.5
Living with a child aged 18 or younger (%)		0.6	0.6	0.6
Living with parents or parent in law (%)		0.7	0.7	0.5
Race	White (%)	85.5	85.6	84.5
	Black (%)	9.2	9.1	9.3
	Asian (%)	2.5	2.5	2.8
	Others (%)	2.8	2.8	3.4
Hispanic (%)		6.0	6.0	6.6
US citizen (%)		98.2	98.2	96.9
Poor English language skills (%)		3.5	3.5	4.0
Educational attainment	Less high school (%)	19.2	19.1	19.8
	High school (%)	31.6	31.8	29.5
	Some college (%)	24.4	24.3	25.6
	BA (%)	13.3	13.4	13.5
	Masters or higher (%)	11.4	11.4	11.6
Income[a]	Mean	$53,749	$54,075	$47,743
	Median	$33,972	$34,240	$28,152
Capital income[a]	Mean	$8852	$8885	$8254
	Median	$0	$0	0
Social Security Income[a]	Mean	$17,088	$17,186	$15,282
	Median	$15,960	$16,050	$14,400
Below poverty threshold (%)		11.3	11.1	15.6
Labor participation (%)		18.2	18.1	16.5
2012 or later (%)		50.7	50.5	53.0
Weighted *n*		338,303,975	320,924,588	17,379,387

Source: American Community Survey, 2005–2017, IPUMS-USA, using *pweight*
[a]Inflation-adjusted 2017 USD, household figures

widowed. This also fits the substantial underrepresentation of married household heads among movers, only 32.7% among movers compared to 44.2% among stayers. However, without a more detailed migration history, this sequencing of events is difficult to establish. Only a small minority of elderly households (12.9%) live with their children, and this is even less so for movers than for stayers.

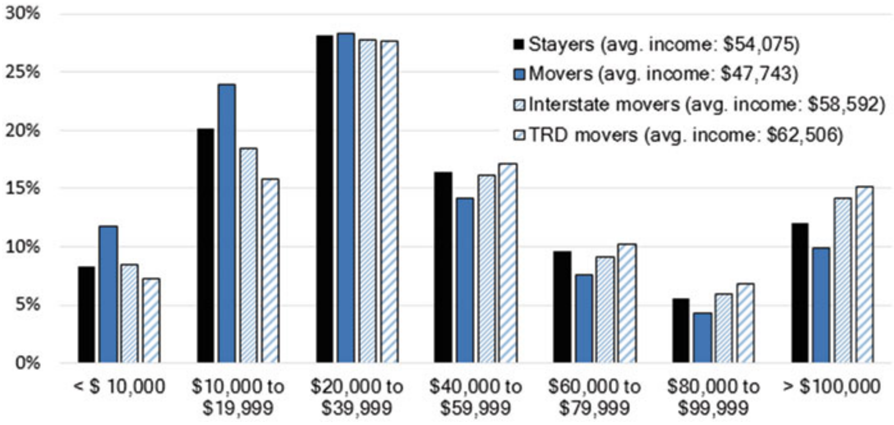


Fig. 12.2 Distribution of household income by migration status. (Source: ACS 2005–2017, IPUMS-USA; TRD = {AZ, FL, GA, NC, SC, TX})

The racial and ethnic composition of the elderly households is predominantly made up of whites, non-Hispanics, and US citizens. This is not surprising as the majority of the more diverse immigrant population entered the USA after the quota system based on natural origin was abolished in the *Immigration and Naturalization Act of 1965*. Not surprising, the movers are slightly more diverse than the stayers, but the differences are small (less than 1 percentage point).

With respect to educational attainment, the differences between movers and stayers are also minimal. Movers are slightly overrepresented, both among the least educated (less than high school degree) and those that have at least some college education. Despite the similarities in educational attainment between movers and stayers, the data do suggest stark income differences. On average, movers earn US$6332 less than stayers. Similar, albeit smaller dissimilarities are also found when just focusing on dividend income and Social Security Income. Moreover, movers are also much more likely to live in poverty. The poverty rate is 15.6% among movers but only 11.1% among stayers. Some of these income and poverty differences may be attributed to the differences in labor force participation: 18.1% of stayers are still working, compared to only 16.5% of movers.

The last variable listed in Table 12.3 is a dummy variable *post-2012* that marks the survey years from 2012 onward. Ideally, one would like to have a time fixed effect to account for macroeconomic conditions throughout the entire study period. However, including age and time fixed effects would not allow us to measure the baby boom cohort effect. The year 2012 was chosen as a divider because it marked the year when the real estate market started its gradual rebound after the housing market crash.

Figure 12.2 provides a first glimpse into the income dependency of late-life migration. The median income bracket for household heads age 65+ ranges from US$20,000 to US$39,999. It includes about 28% of the stayers, as well as 28% of each mover category. In the below-median income brackets, stayers are clearly underrepresented. In the top income brackets, the interstate movers and even more so the TRD movers stand out. Annually, only 5.1% of the elderly households move, and those who move earn a lot less than those who stay. However, among movers, there are substantial income differences by the type of move. The vast majority of movers, 81.4%, stay within the state boundaries, and their income is particularly low, only US$45,271 on average. Current Population Survey (CPS) data for comparable years reveal that only 6.4% of households who moved within their state explicitly name retirement as the main reason to move.

Instead, 50% of them do so for housing-related reasons. About 12% moved for cheaper housing, and, not surprisingly, they constitute the group with the lowest household income. Only 6.4% of the intrastate movers explicitly name retirement as the main reason to move.

The group of interstate movers is smaller, and, on average, interstate movers earn substantially more than stayers and within-state movers. According to CPS data, family reasons motivated migration for about 43% of those households. Less than one fifth of the households named housing-related reasons. Of particular interest are the 12% of the households that indicated retirement as the main reason to move. Their average household income is far above the average, almost identical to the average income of the few households that moved across state boundaries because of a new job or a job transfer. Finally, among interstate movers, 39% moved to one of the traditional retirement destinations (TRD, defined as AZ, FL, GA, NC, SC, and TX). The household income of TRD movers is particularly high, averaging US $62,506. The TRD households that explicitly indicate retirement as the main reason to move make up about 15%. They stand out among the TRD movers, as their average household income is more than a third higher than that of all TRD movers.

Figure 12.3 presents a different perspective on the income variation among the older households. It shows population pyramids that combine three attributes: age, sex, and household income. Separate age-sex-income pyramids are presented for movers (on the right) and stayers (on the left) and for 2005 (top) and 2017 (bottom). The age-sex-income pyramids show that in 2005—and even more so in 2017—young elderly (age 65–74) were overrepresented in the high-income group (US $60,000+). Further, among both movers and stayers, and in all age groups, male household heads are more likely to be in the higher-income group than female household heads. Quite notable is, however, that the income disparity between men and women declined substantially between 2005 and 2017. As argued earlier, women's economic advances are paying off, and they are entering retirement age with more economic power over time. However, in the lowest-income category (under US$30,000), women remain overrepresented, both among movers and among stayers and in all age groups.

We end the data description with a closer look at the spatial characteristics of later-life migration. In absolute terms, the traditional retirement destinations Arizona, Florida, the Carolinas, and Texas play a major role for senior migration. In relative terms, however, Fig. 12.4 shows that senior migration plays a non-negligible role for other states as well. Not surprisingly, the migration flows into three traditional retirement destinations—Florida, Arizona, and South Carolina—are heavily dominated by senior migration. But the same is true for Maine, Arkansas, and several western states, most notable Idaho, Oregon, Nevada, and New Mexico. In contrast, although we categorized Texas as a traditional retirement state and senior

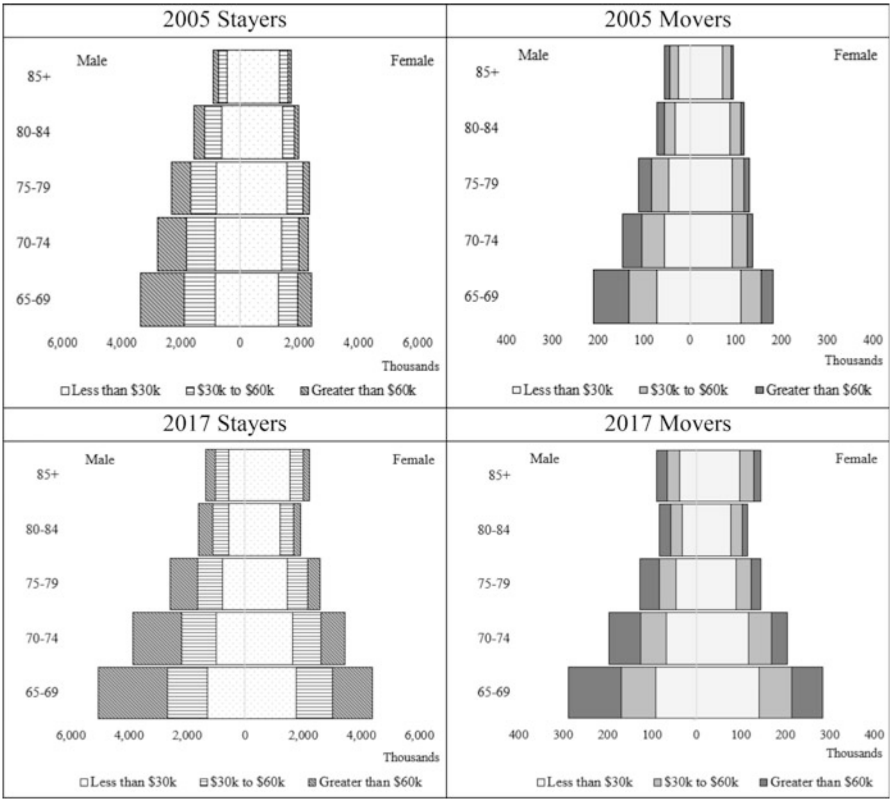


Fig. 12.3 Age-sex-income pyramids for household heads by migration status. (Source: ACS 2005 and 2017, IPUMS-USA)

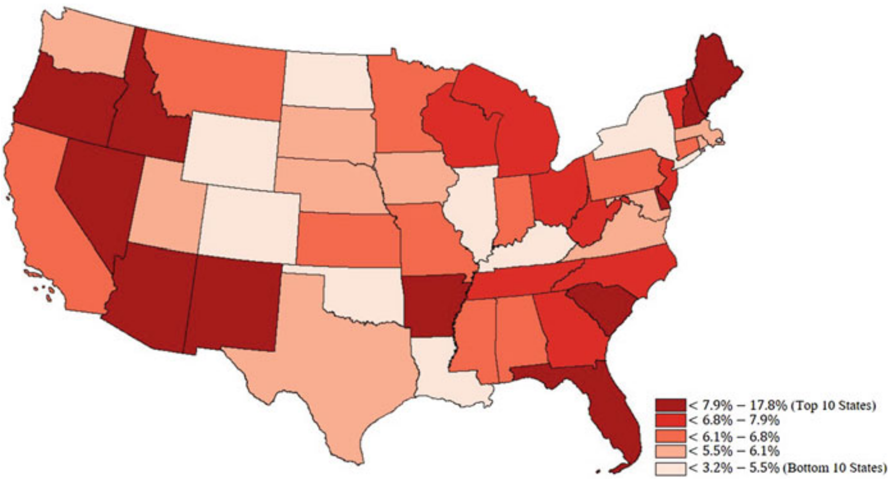


Fig. 12.4 Share of elderly among in-migrants, 2017

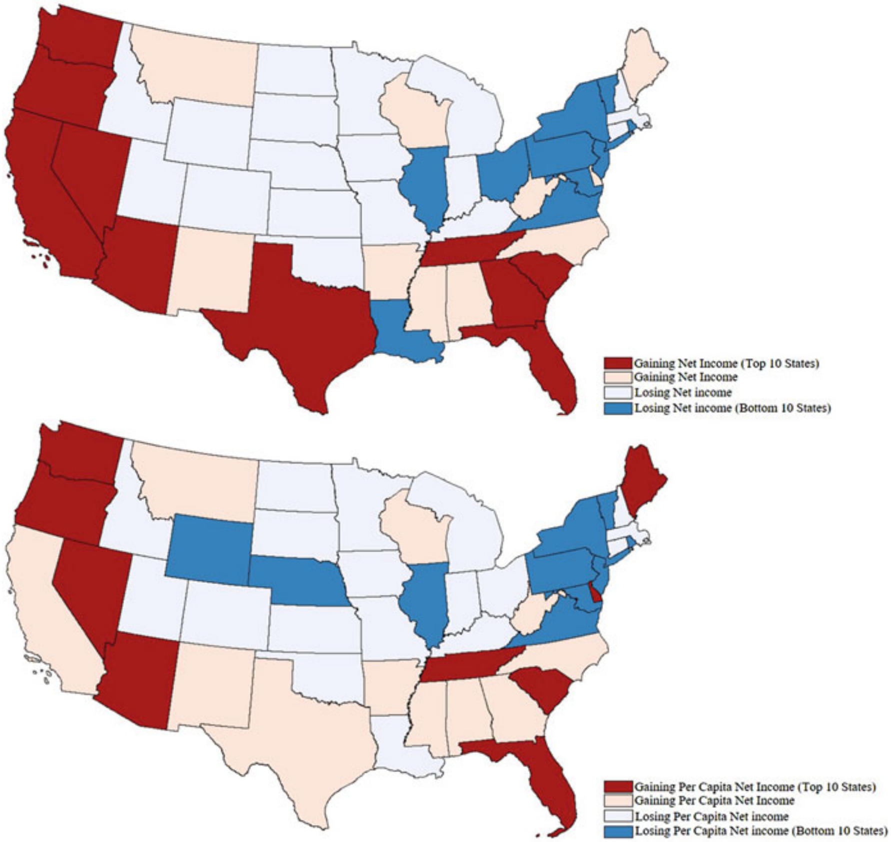


Fig. 12.5 Net income gains and losses due to elderly (65+) migration, 2017. Top, USD; bottom: per capita USD

migration is of importance in absolute terms, it takes on only an average role in relative terms. The reasons are undoubtedly the many economic opportunities in the Houston and Dallas metro areas, which are attractive for younger people still in the labor force. Figure 12.5 provides insight into the spatial income redistribution due to senior migration. The maps show a stark spatial disparity between the states gaining net income and states losing net income. The winners are concentrated along the West Coast and stretch along the southern edge from Arizona to Florida and the Carolinas. The income-losing states are located in the interior, stretching from Idaho to the East Coast and the Canadian border.

12.3.3 Estimation Results

Table 12.4 shows the estimation results for the base models of the three migration variables. The base models include the key variables—income, a dummy for the

Table 12.4 Estimation results: base models

	M_1 Move		M_2 Interstate		M_3 TRD	
	b	SE_b	b	SE_b	b	SE_b
Intercept	−2.730	0.002	−2.219	0.005	−1.693	0.011
Inc (omitted: 40 K–59.9 K)						
<$10,000	0.177	0.001	−0.082	0.004	−0.047	0.008
$10,000–$19,999	0.165	0.001	−0.171	0.002	−0.117	0.004
$20,000–$39,999	0.102	0.001	−0.050	0.002	*0.006*	*0.004*
$60,000–$79,999	−0.074	0.001	−0.026	0.003	0.058	0.005
$80,000–$99,999	−0.077	0.001	0.113	0.003	0.091	0.006
$100,000+	−0.058	0.001	0.151	0.003	−0.023	0.005
Below poverty line	0.159	0.001	−0.100	0.003	−0.025	0.006
Baby boomer	−0.009	0.001	0.062	0.002	−0.016	0.004
Female	−0.095	0.001	−0.026	0.001	−0.152	0.003
Control variables						
With spouse	−0.434	0.001	0.207	0.002	0.262	0.003
Age (omitted: 85+)						
Ages 65–69	0.13	0.001	0.462	0.002	0.352	0.005
Age 70–74	−0.034	0.001	0.335	0.002	0.401	0.004
Age 75–79	−0.118	0.001	0.236	0.002	0.284	0.005
Age 80–84	−0.127	0.001	0.098	0.003	0.076	0.005
Race (omitted: other race)						
White	−0.100	0.001	0.166	0.004	0.699	0.009
Black	−0.210	0.002	−0.220	0.005	0.885	0.010
Asian	*−0.003*	*0.002*	−0.150	0.006	0.142	0.012
Hispanic	−0.009	0.001	−0.176	0.003	0.604	0.006
Not US citizen	0.523	0.002	−0.548	0.005	0.267	0.010
Poor English	−0.139	0.002	−0.275	0.005	−0.184	0.010
Educ (omitted: < HS)						
High school	−0.018	0.001	0.197	0.002	0.154	0.004
Some college	0.150	0.001	0.409	0.002	0.180	0.004
Bachelor's	0.173	0.001	0.555	0.002	0.140	0.005
Masters or higher	0.222	0.001	0.707	0.003	−0.010	0.005
Labor force	−0.109	0.001	−0.533	0.002	−0.253	0.004
Capital income	0.006	0.000	−0.005	0.000	0.005	0.000
Social Security Inc.	−0.042	0.000	0.069	0.001	0.063	0.001
2012 or later	0.090	0.001	−0.045	0.001	0.077	0.003
n	4,172,204		188,858		35,550	
Weighted *n*	338,303,975		17,379,387		3,224,623	
$-2\text{Log}L$	137,037,246		16,673,620		4,315,011	

Note: estimates in *italics* are not significantly different from zero at $\alpha = 5\%$

baby boom cohort, and a dummy for female household heads—and the basic controls. But they exclude fixed effects and interaction terms.

To begin the presentation of the estimation results, we briefly turn to the estimated parameters for the control variables. Married-couple households are substantially less likely to move than household heads without a spouse. However, if they move, their destination is more likely to be a different state and especially a traditional retirement state. The estimated impact of age on the moving behavior of the elderly population is, by and large, in line with previous studies. First, the probability of moving, $\Pr(M_1 = 1)$, declines with increasing age, for all, but the oldest age cohort (85+). Second, among those who do make a move, the negative age dependency of the probability that the move is an interstate move, $\Pr(M_2 = 1)$, is even stronger and does not include the uptick for the oldest old. For example, compared to the oldest old, the odds of the 65–69-year-old movers crossing state boundaries are 1.6 (=exp (0.462)) times higher. They are 1.4 times higher for the 70–74-year-olds and 1.27 times higher for the 75–79-year-olds. Lastly, the probability that an interstate migrant relocates to a traditional retirement destination, $\Pr(M_3 = 1)$, peaks for those in the early 70s and rapidly drops at older ages. Compared to the oldest old interstate movers, the odds of choosing a TRD is 1.4 times higher for a 65–69-year-old and 1.5 times higher for a 70–74-year-old. This may very well reflect the often reported migration of the oldest old to locations in close proximity to their children.

With respect to estimated racial and ethnic variations, black elderly are the least likely to move and the least likely to make an interstate move. However, if they do make an interstate move, they are the most likely to move to a traditional retirement destination. This may very well reflect the historical position and trajectory of blacks in the US South. Similarly, elderly Hispanics are also less likely to move, less likely to move across state boundaries, and more prone to choose a traditional retirement destination. This pattern is expected given that the traditional retirement destinations—especially Arizona, Florida, and Texas—include large Hispanic population clusters. For the most part, however, elderly migration patterns are very much dominated by whites. Whites make up 85% of the elderly households, and they are by far the most likely to migrate across state boundaries.

Differential moving behavior by educational attainment level reveals some unexpected patterns. The propensity to move as well as the propensity to make an interstate move is rising with rising educational attainment levels. Yet, moving to a traditional retirement state is, ceteris paribus, only weakly related to education. In fact, highly educated movers, i.e., those with at least a master's degree, are the least likely to choose a traditional retirement destination.

Turning now to the important income dependency of elderly migration behavior, recall that we hypothesized that the affluent elderly are more likely to move, more likely to cross state boundaries, and more likely to choose a traditional retirement destination. The results of the base model—the estimated odds ratios relative to the omitted income category ($40,000–$59,999) are shown in Fig. 12.6—suggest a more nuanced picture. Most notably, we find that the affluent are actually quite satisfied with their residential situation, as they have a low migration propensity. The poverty variable, which takes income and household size into account, confirms

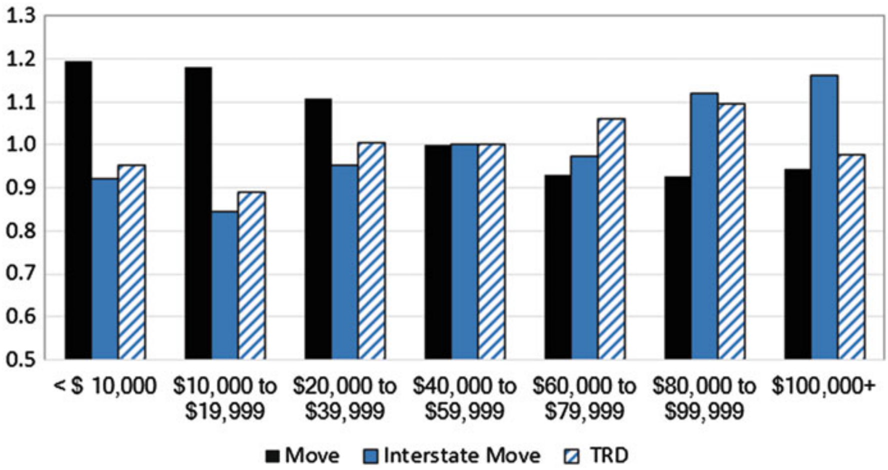


Fig. 12.6 Estimated relative odds derived from base model

these results. The inverse link between income and moving propensity may very well be connected to homeownership. That is, many low-income households are renters rather than homeowners, and renters are much more mobile than homeowners.[4]

As hypothesized, movers' choice to relocate to another state becomes more likely with rising income. The estimated parameters are significant but are not substantial. For example, the odds that a migrant belonging to the highest-income group chooses a destination in another state is only 1.16 times higher than that of comparable migrants from the middle-income group. For the third migration variable—the probability that interstate movers select one of the traditional retirement destination states—the base model suggests a significantly positive but extremely weak association with income.

With respect to the two demographic markers—cohort and sex—the base model suggests that baby boomers, by and large, behave similarly as earlier cohorts: similar moving propensity and similar selection probability of traditional retirement destination. Baby boomers are, however, slightly more prone to making interstate moves. However, while the difference is significantly different from zero, its magnitude is negligible. For women, we find that—compared to men—they are less likely to move, slightly but not significantly less likely to move across state boundaries and less likely to head toward a traditional retirement destination.

It is possible that the estimated income, cohort, and gender differences are weak or even hidden because the models are aspatial and ignore, for example, that retirement migration patterns could be structurally different for those originating in the Midwest compared to those originating in the South. If such spatial migration

[4]Unfortunately, the ACS does not provide information on homeownership status prior to the move.

Table 12.5 Base model and origin FE model in comparison: odds ratios

	M_1 Move		M_2 Interstate		M_3 TRD	
	Base	FE	Base	FE	Base	FE
<$10,000	1.19	1.39	0.92	0.61	0.95	0.83
$10,000–$19,999	1.18	1.37	0.84	0.60	0.89	0.82
$20,000–$39,999	1.11	1.24	0.95	0.72	1.01	0.95
$40,000–$59,999	1.00	1.00	1.00	1.00	1.00	1.00
$60,000–$79,999	0.93	1.00	0.97	0.81	1.06	1.01
$80,000–$99,999	0.93	0.99	1.12	0.95	1.10	1.03
$100,000+	0.94	1.01	1.16	1.01	0.98	0.91
< Poverty line	1.17	1.18	0.90	0.89	0.97	*1.00*
Female	0.91	0.96	0.97	0.88	0.86	0.82
Baby boomer	0.99	0.98	1.06	1.14	0.98	*1.01*

Note: Estimates in *italics* are not significantly different from 1 at $\alpha = 5\%$

regimes exist, then the models need to control for migrants' origins. We did this by adding origin state fixed effects. Table 12.5 juxtaposes the estimated odds ratios of the base model and the origin fixed effect (FE) model. For the estimated income dependency, the fixed effect model suggests a slightly different interpretation than the base model. For the low-income groups, the propensity to move ($M_1 = 1$) is higher than the base model suggests, and it does not show any variation among those making US$40,000 or more. Further, for the probabilities to make an interstate move ($M_2 = 1$) or a TRD-move ($M_3 = 1$), the FE models suggest that they increase with rising income only for those with incomes below US$40,000. The results of the base model and the origin state FE model are comparable with respect to the estimated generational and gender effects.

The estimated parameters of the fixed effect models also provide a glimpse into the spatial regimes of elderly migration. Figure 12.7 shows the estimated odds ratios of the origin fixed effects of the three migration models. The first map (top) suggests a strong East-West disparity. Seniors living in the eastern USA are more likely to be stayers than those living in the west of the Mississippi. Exceptions are seniors living in Florida and, to a lesser extent, in Illinois and Maine. The estimated odds ratios of the origin fixed effects for making an interstate move suggest a more complex geography. Low propensities to make interstate moves are prevalent along the West Coast, as well as in a broad interior area that stretches from Midwestern states along the Canadian border to Texas. In between are the mountain states and the Dakotas where elderly migrants have an above-average propensity to cross state borders. The same is true for most states along the eastern seaboard. The map at the bottom of Fig. 12.7 suggests that interstate migrants' decisions to relocate to a traditional retirement destination also show an East-West disparity. Not surprisingly, some of the highest propensities to move to a traditional retirement destination are found in states with a cold winter climate, from Minnesota to Maine.

To properly address the hypothesis of differential income/wealth dependencies across generations, we confine the sample to persons of age 65–71. Without this restriction, baby boomers would be heavily underrepresented. We further use

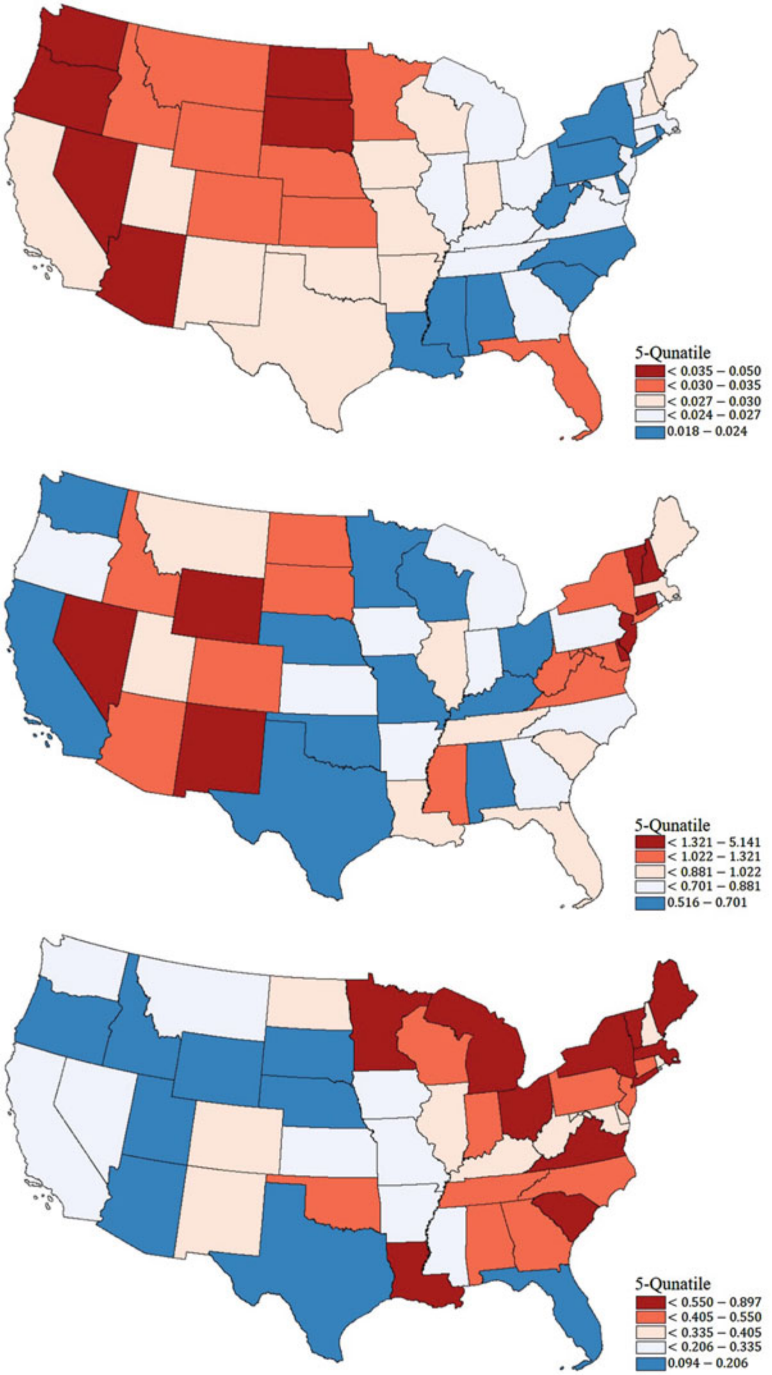


Fig. 12.7 Quantile maps for estimated FE odds ratios of origin FE models of M_1 (top), M_2 (middle), and M_3 (bottom)

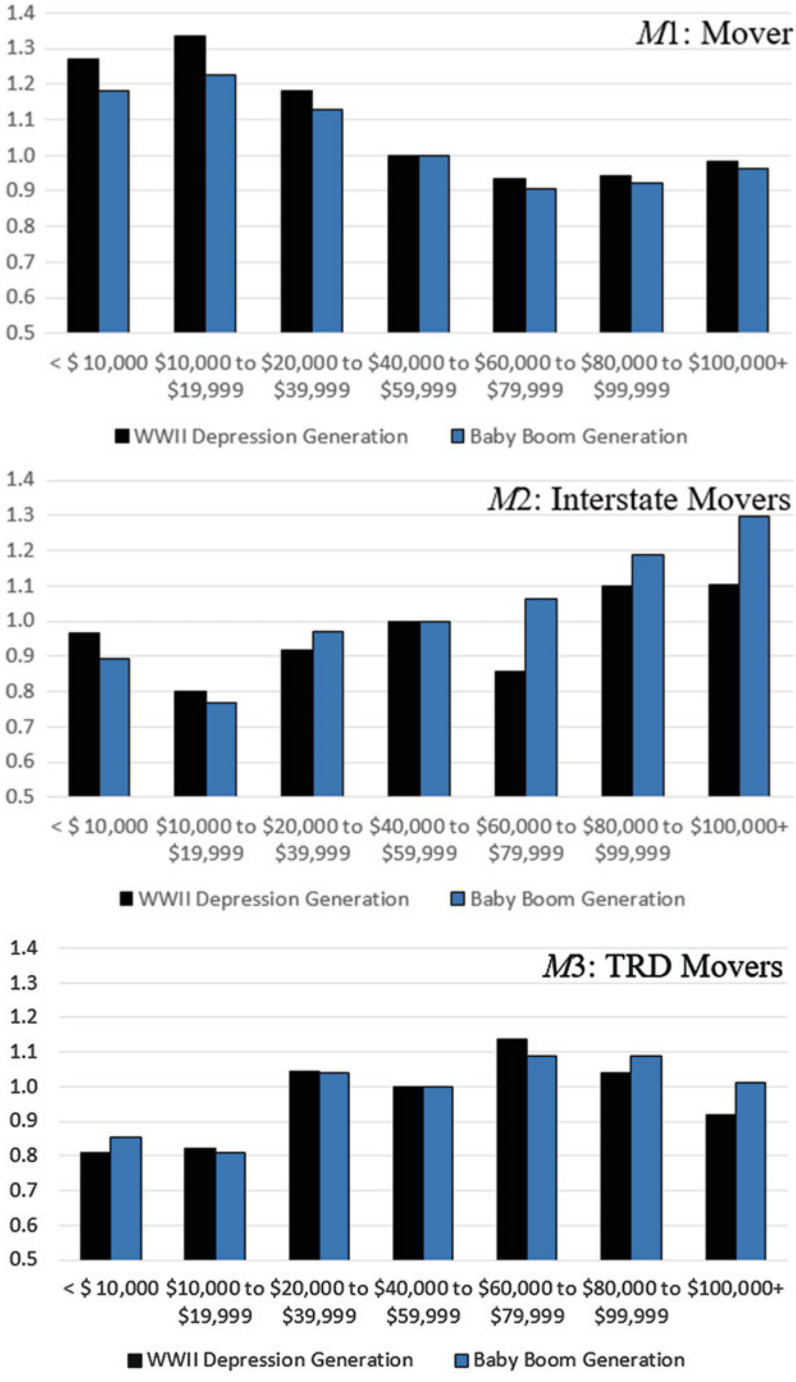


Fig. 12.8 Estimated relative odds of moving by income group, baby boom cohort versus WWII cohort, age 65–71, during 2005–2017

individual ages rather than age groups, and, most importantly, we enhance the FE model with the set of interaction variables X∗Boomer. Figure 12.8 shows the key results in the form of income odds ratios by cohort. The moving propensities of both baby boomers and its predecessor WWII generation exhibit an income decay. However, the steepness of the decay is significantly less pronounced for the baby boomers than for the WWII generation. For the most part, the generational differences are not substantial, and for the upper income groups, the decay flattens out for both cohorts. The two generations are more dissimilar with respect to movers' propensity to cross state boundaries ($M_2 = 1$). We find that the relative odds of crossing state borders systematically rise with income. For the WWII generation, we cannot find such a clear pattern. Lastly, we also do not find a systematic pattern for the income dependency of choosing a retirement destination for either of the two generations.

The last issue deals with gender differences in income dependency of elderly migration behavior. When modifying the FE model with the set of interaction variables X∗Female, we do not find stark gender differences for moving versus staying and for the probabilities of crossing state lines. Only for the third migration variable, moving to a traditional retirement destination, do we find gender differences. Specifically, we find that men's probabilities to move to one of the retirement states rises with income. But for women, we cannot discern a distinct association with income. Since only 32.5% of the men compared to 80.7% of the women are moving without a spouse, it may very well be that the differences are entangled in the interaction of gender and marital status differences.

12.4 Conclusion

In 2010, the first members of the baby boom generation reached retirement age. Many of them follow the example of their predecessor generation and move to a different location away from work, often in combination with downsizing. This research focuses on how the nexus of aging, migration, and income plays out for the older population, with a special emphasis on the baby boom generation. Employing the 2005–2017 ACS samples of the US population age 65 and older, we find that about 5% of that population moves every year. Almost 19% of those moves cross state boundaries, of which two-fifth are headed to a traditional retirement destination. Stayers, on average, earn more than those who only make residential adjustments within the state. The average income is much higher for interstate movers and in particular for those who move to a traditional retirement state. This results in a spatial income redistribution with a distinct North-South disparity. We find that—in terms of net income gains—the winners of retirement migration are states located in the West and South, whereas the interior and the East Coast are losing net income.

Older person's migration across interstate boundaries creates this stark income gain disparity because it is a highly selective process. We find that those choosing to move out of state are predominantly white, affluent, and better educated.

Importantly, we also find a strong positive association between income and interstate moves for the baby boomers, but not for the WWII generation. Given that baby boomers are slightly overrepresented, it is reasonable to expect that this income selectivity, and the associated spatial income redistribution, will become stronger in the future when baby boomers become more and more dominant among the older population.

Finally, we do recommend future research on women's retirement moving behavior. Women's high life expectancy and their growing economic power within the baby boom generation are likely to contribute to migration differences between baby boomers and their predecessors. We only found weak gender differences and suspect that gender alone is not the driving force that defines possible differences between men's and women's migration behavior. Instead, we suggest to further investigate how the various manifestations of the marital status—gender—income/wealth complex affect migration behavior.

References

Case A, Deaton A (2015) Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21st century. Proc Natl Acad Sci 112(49):15078–15083

Clark DE, Knapp TA, White NE (1996) Personal and location-specific characteristics and elderly interstate migration. Growth Chang 27(3):327–351

Conway KS, Houtenville AJ (2001) Elderly migration and state fiscal policy: evidence from the 1990 census migration flows. Natl Tax J 54(1):103–123

Conway KS, Houtenville AJ (2003) Out with the old, in with the old: a closer look at younger versus older elderly migration. Soc Sci Q 84(2):309–328

Davies A (2014) Urban to rural elderly migration: renewing and reinventing Australia's small rural towns. In: Dufty-Jones R, Connell J (eds) Rural change in Australia. Ashgate, Farnham, pp 43–55

Day FA, Barlett JM (2000) Economic impact of retirement migration on the Texas Hill Country. J Appl Gerontol 19(1):78–94

Duncombe W, Robbins M, Wolf DA (2001) Retire to where? A discrete choice model of residential location. Int J Popul Geogr 7(4):281–293

Graves P, Linneman P (1979) Household migration: theoretical and empirical results. J Urban Econ 6(3):383–404

Haas WH III, Serow WJ (2002) The baby boom, amenity retirement migration, and retirement communities: will the golden age of retirement continue? Res Aging 24(1):150–164

Kromer B, Howard D (2013) Labor force participation and work status of people 65 years and older. American Community Survey Briefs 11-09, US Census Bureau. http://www.census.gov/prod/2013pubs/acsbr11-09.pdf. Accessed Nov 2018

Longino CF Jr, Bradley DE (2003) A first look at retirement migration trends in 2000. The Gerontologist 43(6):904–907

Plane DA, Heins F (2003) Age articulation of US inter-metropolitan migration flows. Ann Reg Sci 37(1):107–130

Plane DA, Jurjevich JR (2009) Ties that no longer bind? The patterns and repercussions of age-articulated migration. Prof Geogr 61(1):4–2

Rogers A (1988) Age patterns of elderly migration: an international comparison. Demography 25 (3):355–370

Rowles GD, Watkins JF (1993) Elderly migration and development in small communities. Growth Chang 24(4):509–538

Ruggles S, Flood S, Goeken R, Grover J, Meyer E, Pacas J, Sobek M (2018) IPUMS USA: Version 8.0 [dataset]. IPUMS, Minneapolis. https://doi.org/10.18128/D010.V8.0. Accessed October 2018

Serow WJ (2001) Retirement migration counties in the Southeastern United States: geographic, demographic, and economic correlates. The Gerontologist 41(2):220–228

Serow WJ (2003) Economic consequences of retiree concentrations: a review of North American studies. The Gerontologist 43(6):897–903

Sewdas R, de Wind A, van der Zwaan LG, van der Borg WE, Steenbeek R, van der Beek AJ, Boot CR (2017) Why older workers work beyond the retirement age: a qualitative study. BMC Public Health 17(1):1–9

Smith SK, House M (2006) Snowbirds, sunbirds, and stayers: seasonal migration of elderly adults in Florida. J Gerontol Ser B 61(5):S232–S239

Somenahalli S, Shipton M (2013) Examining the distribution of the elderly and accessibility to essential services. Procedia Soc Behav Sci 104:942–951

U.S. Census Bureau (2018) Older people projected to outnumber children for first time in U.S. history. Newsroom. https://www.census.gov/newsroom/press-releases/2018/cb18-41-population-projections.html. Accessed Nov 2018

Waldorf BS, McKendree M (2013) The aging of rural Indiana's population. Rural Indiana Issues # 4 EC-769-W. http://www.agecon.purdue.edu/extension/pubs/files/Aging_EC-769-W.pdf. Accessed Nov 2018

Waldorf BS, Pitfield D (2005) The effect of demographic shifts on the transportation needs of the elderly. In: Donaghy K, Poppelreuter S, Rudinger G (eds) Social dimensions of sustainable transport: transatlantic perspectives. Ashgate, Aldershot, pp 67–89

Walters WH (2000) Assessing the impact of place characteristics on human migration. Area 32 (1):119–123

Walters WH (2002) Later-life migration in the United States: a review of recent research. J Plan Lit 17(1):37–66

Whisler RL, Waldorf BS, Mulligan GF, Plane DA (2008) Quality of life and the migration of the college-educated: a life-course approach. Growth Chang 39(1):58–94

Wiseman RF, Roseman CC (1979) A typology of elderly migration based on the decision making process. Econ Geogr 55(4):324–337

Chapter 13
The Demography of Water Use: Why the Past Is a Poor Predictor of the Future

Patricia Gober

Abstract The concept of gallons per capita per day (GPCD) expresses the relationship between population and water. US trends in GPCD between 1950 and 2015 reflect changing technology, policy, and lifestyles. Water managers use GPCD for long-term infrastructure planning and water resource management. GPCD grew from 1950 to 1980, remained steady from 1980 to 2000, and then fell sharply due to the implementation of the 1992 US Energy Policy Act. This legislation mandated water-efficient fixtures and appliances in new and remodeled homes and focused on indoor water use. The new frontier of water conservation is now outdoor use where behavioral and cultural forces interact with policy and technology to reduce GPCD. Regional trends reveal significantly higher water use in western than eastern states due in part to their warm, dry climates. California, Texas, Nevada, and New Mexico led a regional decline with some other western states unaffected by national trends. The rapidly changing and geographically dispersed pattern of declining GPCD has refocused attention to outdoor water. Uncertainties associated with climate change; city, state, and federal conservation policies; lifestyles; and public attitudes imply a change in water management practices from traditional predict-and-plan methods of long-term water planning to scenario planning, exploratory modeling, and decision-making under uncertainty (DMUU).

Keywords GPCD (gallons per capita per day) · Water conservation · Outdoor use · Climate change · Uncertainty

---

P. Gober (✉)
School of Geographical Sciences and Urban Planning, Arizona State University, Tempe, AZ, USA
e-mail: gober@asu.edu

R. S. Franklin (ed.), *Population, Place, and Spatial Interaction*, New Frontiers in Regional Science: Asian Perspectives 40,
https://doi.org/10.1007/978-981-13-9231-3_13

13.1 Introduction

This chapter tracks the recent history of gallons per capita per day (GPCD) in the USA and discusses what it means for future planning and policy. Until recently, urban water managers treated population growth as the major determinant of future water demand and the basis for revenue projections and infrastructure capacity planning. They applied a straight-line projection, assuming that gallons per capita per day (GPCD) would remain constant, paying little attention to other determinants of water demand (Quay 2015). Since 1990, many urban communities experienced declining GPCD, averaging 0.5% per year (Coomes et al. 2010). Water managers, caught by surprise to discover that their assumption about constant GPCD was inaccurate, began to look beyond population growth to climate, technology, policy, and lifestyles for insight into future trends in water demand (Gober et al. 2016). Also significant are the uncertainties associated with these issues for traditional predict-and-plan methods of long-term water planning.

13.2 National Trends in Water Use

Trends in water withdrawals at the national level reflect basic changes in the relationship between population and water. The United States Geological Survey (USGS) has monitored water use in the nation and its constituent states at five-year intervals since 1950. Public water supplies refer to water withdrawn by public suppliers for domestic, commercial, and industrial purposes. Public suppliers refer to agencies that provide water to at least 25 people. Public supply water supports domestic, commercial, and industrial purposes. Also included are public services such as firefighting and system losses (leakages) (Deiter et al. 2019).

Water withdrawals (from surface and groundwater sources) for public supply tracked population closely until 1980 when American lifestyles favored larger homes, more outdoor irrigation, backyard pools and spas, and high water-using fixtures and appliances (Fig. 13.1). Water withdrawals grew faster than population between 1980 and 2005. Prevailing lifestyles required higher GPCD. After 2005, total water withdrawals fell significantly from 44.2 in 2005 to 39 billion gallons per day in 2015, the equivalent of 11.7% (Deiter et al. 2019). This remarkable turn-around meant that lower water withdrawals supported a growing national population.

Trends in GPCD echoed this larger story about population and water. GPCD grew steadily from 145 GPCD in 1950 to 183 in 1980. High use levels pertained through the end of the twentieth century and then fell to 157 in 2010 and 138 in 2015 (Fig. 13.2). Declining per capita use is not explained by changes in household structure as average household size steadily declined during this period from 2.62 in 1990 to 2.54 in 2015 (US Census 2019), and average household size is negatively related to per capita use. Smaller households are associated with higher per capita use because they are unable to take advantage of the economies of scale at the

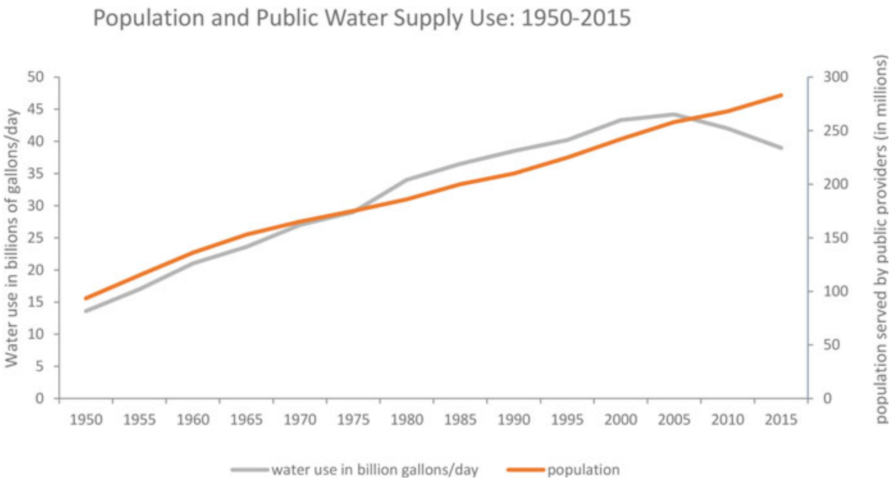


Fig. 13.1 Population and water use in the USA, 1950–2015. (Source: US Geological Survey. Estimated Use of Water in the United States in 1950–2015)

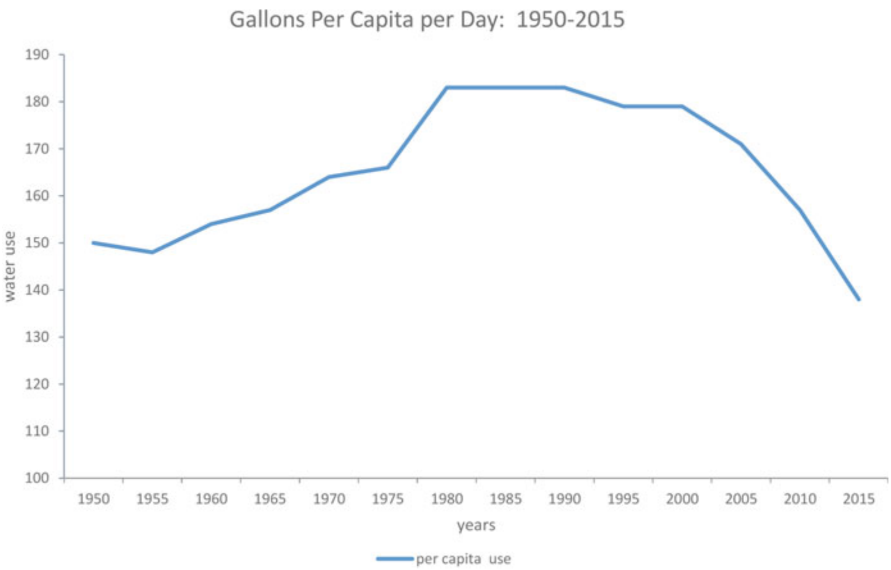


Fig. 13.2 Per capita water use in the USA, 1950–2015. (Source: US Geological Survey. Estimated Use of Water in the United States in 1950–2015)

household level (Höglund 1999; Arbués et al. 2010). Larger households spread fixed outdoor water use and indoor use associated with appliances such as dishwashers and clothes washers over a larger number of people. Much of the decline in GPCD between 2005 and 2015 occurred despite, not because of, the trend toward smaller households.

Using water utility records from Louisville, Kentucky, Coomes et al. (2010) found that homes built after 1994 used about 13 gallons per day less than homes built before 1994, controlling for size and value. They concluded that the introduction of low-flow toilets, showers, and clothes washers had a significant effect on residential water use, accounting for a decline of about 16% on average over 20 years. While the changes from 1 year to the next were relatively small, the cumulative effects of new fixtures and appliances reduced household water use by 16% in 20 years.

Indoor water conservation was largely a policy-driven process (Vickers and Bracciano 2014). The 1992 US Energy Policy Act, implemented in 1994, established national water efficiency requirements and set maximum flow rates for all plumbing fixtures installed in new and renovated homes. The Environmental Protection Agency's WaterSense Program launched in 2001 added third-party certification for water fixtures and appliances with more stringent but voluntary standards. These requirements were mandated by many communities and became the industry standard for new construction and retrofits. Declining GPCD, illustrated by USGS data, reflected the gradual penetration of low-volume fixtures and appliances in North American homes (Coomes et al. 2010; Kiefer et al. 2013).

It is unclear whether the same policy-driven process that reduced indoor water use will be equally successful for reducing outdoor use. Outdoor water use is climate sensitive and related to the host of policy, social, behavioral, and lifestyle issues. Outdoor water use goes to the heart of why and how people landscape their homes with irrigated plants, use pools and spas, maintain domestic irrigation systems, hire outside contractors to manage their lawns and gardens, and adhere to conservation directives. Reducing outdoor water use involves a far more complex set of environmental and social issues than the technologies and policies that led to decline in indoor water use.

13.3 Regional Trends in Water Use

While the national story of declining per capita water use signals the start of a new relationship between population and water use, the regional picture is more complicated. There was a clear East/West divide in per capita water use in 2005 (Fig. 13.3). Domestic per capita water rates were above 100 GPCD in the West and below that mark in most eastern states. The highest public supply use rates were in Texas (197), Idaho (180), Nevada (189), Utah (186), and Wyoming (167), with the entire West blanketed by high per capita use. In contrast, most states in the East and Midwest used significantly less water. The lowest use rates occurred in Maine (51), Pennsylvania (56), Ohio (68), and New Hampshire (75). Deiter et al. 2019) attribute the East/West divide to higher levels of outdoor water use for irrigation in the warm, dry climate of the West.

The GPCD use pattern became more complex by 2015 as some western states significantly reduced use rates between 2005 and 2015 (Figs. 13.3b and 13.4). Four

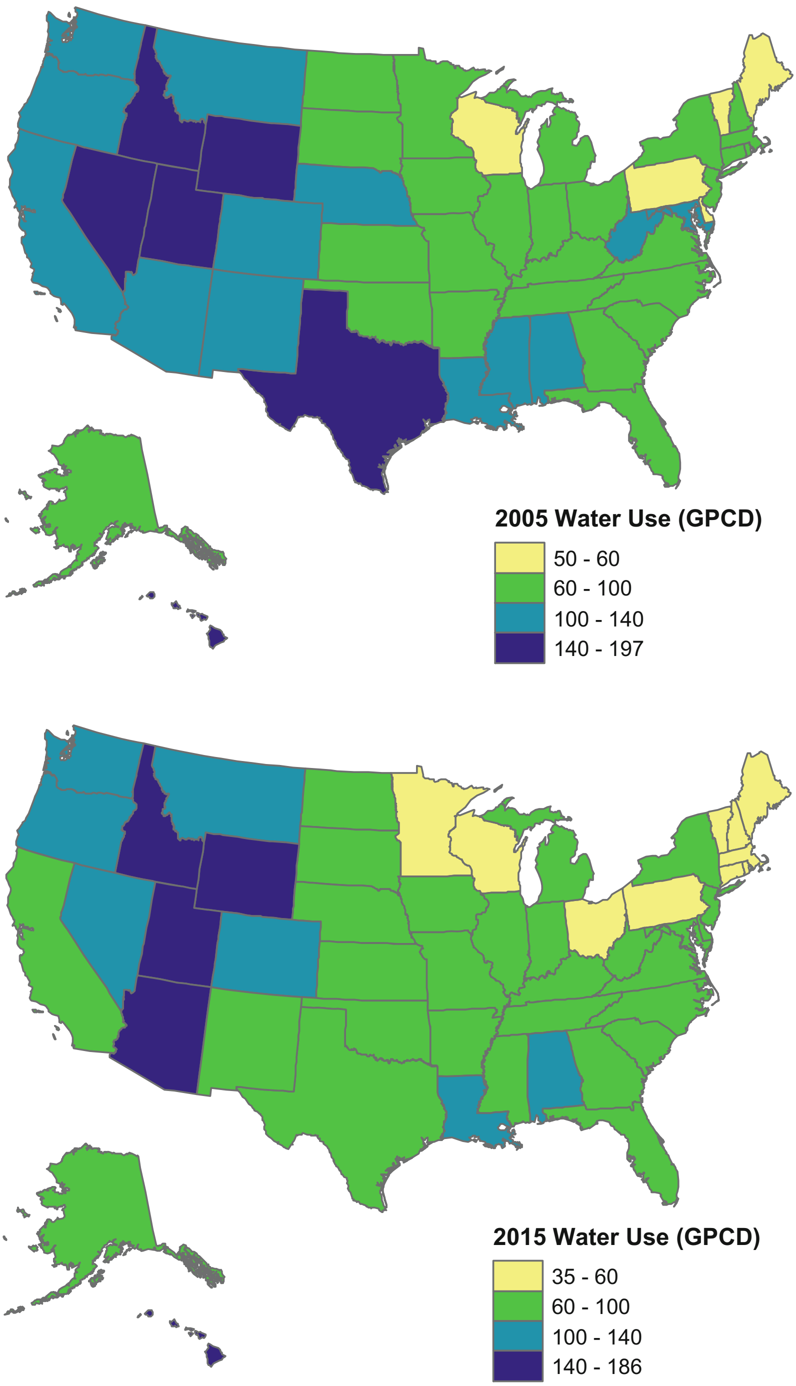


Fig. 13.3 (**a** and **b**) GPCD by state, (**a**) 2005 and (**b**) 2015. (Source: US Geological Survey. Estimated Use of Water in the United States in 1950–2015)

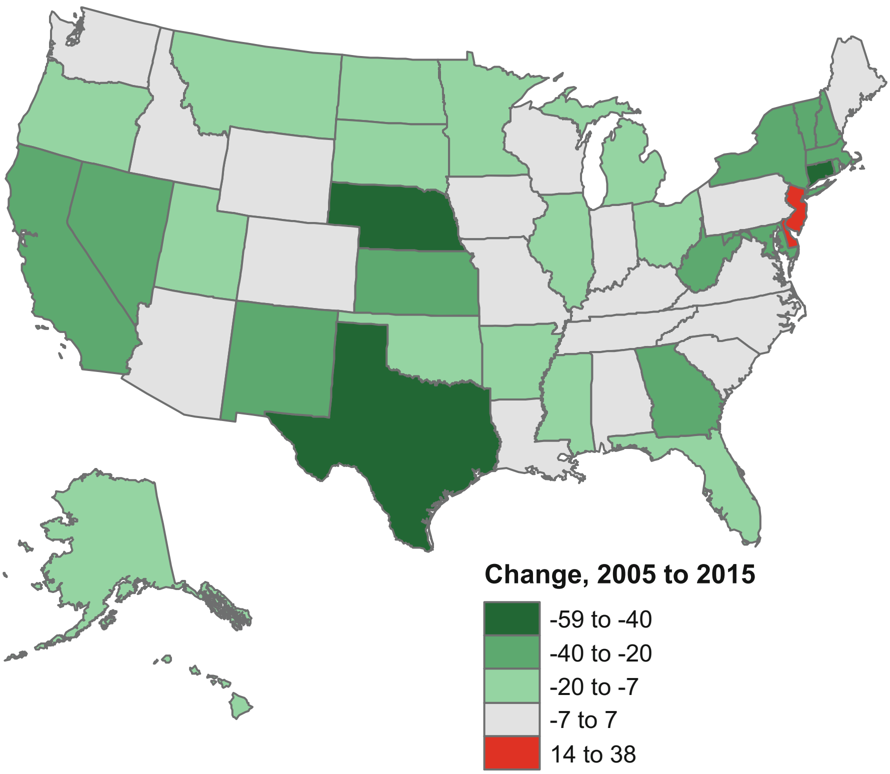


Fig. 13.4 Percent change in GPCD, 2005–2015. (Source: US Geological Survey. Estimated Use of Water in the United States in 1950–2015)

key examples were Texas (from 197 in 2005 to 81 in 2015), New Mexico (113–87), California (119 to 86), and Nevada (189–122). The four states experienced significant challenges to their water supplies, including drought in California, low water levels in Lake Mead on the Colorado River in Nevada, declining water levels in the Ogallala Aquifer in Texas where it represents 40% of the water supply, and severe declines in groundwater levels in parts of New Mexico. Together, they represented 27.7% of the nation's public water supply withdrawals in 2005, declining to 21.3% in 2015. The decline in per capita water use in the West was not so much a sweeping regional phenomenon but one anchored in key states experiencing severe environmental stress coupled with rapid population growth. Water conservation became an important new supply source and vehicle to support continued population growth. The Southern Nevada Water Authority's most recent plan assumes a decrease GPCD from 127 in 2017 to 116 in 2035 and 111 in 2050 (Southern Nevada Water Authority 2018). These reductions are in use and reuse of indoor water figure into supply and demand scenarios for future water resources planning. The goal is to balance population growth with water conservation to achieve sustainable water use in the future.

California's most recent drought experience illustrated the new relationship between population and water in the West. The state, like most of the western states, owes its modern development to federally funded dams, reservoirs, and canals (Reisner 1986). These projects initially supported irrigated agriculture in downstream valleys and today support large urban populations, agriculture, and environmental flows. AghaKouchak et al. (2015) argued that California reached the limits of building infrastructure to boost supply as population and water demand grew. The most recent drought episode demonstrated the folly of overuse and poor management. The event occurred between fall 2011 and early 2016. It was the driest period since record keeping began in 1895. 2014 and 2015 were the two hottest years in the state's history (Hanak et al. 2016). Governor Edmund Gerald Brown Jr. declared a statewide drought emergency in April 2015 and ordered a 25% reduction (relative to 2013) in water use for cities and towns. Cities were able to cope as per capita water use declined sharply. Consequences for agriculture and the environment were more severe. Growers in rural areas received 50% less irrigation water. Extinction threatened 18 fish species, wildlife refuges experienced shortages, and wildfire risk increased in many areas.

Recognizing that this was not just another one-off event, Brown announced in 2016 that the water restrictions put into place during the drought event would be permanent and challenged the state to adjust to a "new normal." "Ongoing drought conditions and our changing climate require California to move beyond temporary drought measures and adopt permanent measures to use water more wisely and to prepare for more frequent and persistent periods of limited water supply." (Executive Department State of California 2016). The California experience was symbolic of the need for a new relationship between water and development in the West, a recognition that sustainable growth would require lower GPCD, difficult trade-offs between agriculture, cities, and the environment, and public policy would play an increasing important role in the relationship between population and water. California's per capita water use fell by more than 30% during this tumultuous period between 2010 and 2015.

Policy and technology played an important role in improving indoor water efficiencies. As the need for future gains in efficiencies shift to outdoor water, it is less clear that policy solutions alone can deliver the requisite declines in GPCD. Outdoor water use implicates lifestyles, identity, cultural preferences, and economic imperatives more than indoor use. Policy research has emphasized the effects of price on behavior, with far less insight into the potential effects of mandatory and voluntary drought restrictions on water use. Climate change also affects outdoor water use, translating long-term water planning from a predict-and-plan problem to one of deep uncertainties requiring new strategies for decision-making and modeling.

13.4 Outdoor Water Use: What Do We Know?

Empirical evidence indicates that homeowners and their landscaping contractors do not always efficiently manage their landscape designs. Moreover, many of these designs are ill suited to local weather and climate conditions. In a study of 700 single-family homes across California water agencies, De Oreo et al. (2011) found that 87% of homes irrigated their yards, 54% of homes that irrigate did so in excess, and 62% of excess use occurred on just 15% of lots. Results imply there are cultural and behavioral dimensions to outdoor water use and that technology and policy solutions may not be enough to reduce the use. They also suggest that targeted water conservation and educational campaigns and enforcement activities are required to target the relatively small number of households that over-irrigate.

The field of landscape architecture offers insight into why heavily watered landscapes were adapted in the US West where they were often ill suited to local environmental conditions. In the *Crabgrass Frontier*, Kenneth Jackson described a nineteenth-century society of middle-class Americans who sought after "pleasure gardens" as a sign of suburban respectability (Jackson 1985). They saw the gardens surrounding their homes as good places to raise children and a mark of suburban respectability. With westward settlement, these idealized gardens transferred to climates where they required high inputs of artificial fertilizers and irrigation water. In *Lawn People*, Robbins (2007) described the role of pride, status, identity, and political economy to explain the preference for heavily watered gardens. Green lawns and gardens are not only individual desires but also a manifestation of social status and need for collective identity. They are part of a global lawn-care industry that markets fertilizers and seeds, lawn mowers, and irrigation equipment.

Perceptions and attitudes play a role in how people use outdoor water. Russell and Fielding (2010) attribute conservation behavior to attitudes (how positively or negatively particular policies are viewed), beliefs (worldview about the relationship between humans and the natural world), routines and habits (stable behavior patterns that have been reinforced in the past), personal capabilities (knowledge and skills to implement conservation practices), and contextual factors (e.g., policies, incentives, price). Neel et al. (2014) investigated the psychological aspects of high water-use residential landscapes in an arid southwestern city. They found that that the subjects in their study associated high water-use landscapes with sexual attractiveness and family orientation. Thus, if a person wants to convey sexual attractiveness and family status, they would favor landscapes with grass trees and other high water-use plants.

Bollinger et al. (2018) studied peer effects in the use of water conservation in terms of the diffusion of dry landscaping. In a study of water use in Phoenix, they found there was a much higher probability of a switch from green to dry landscaping just after a house transaction. They investigated the effects of these changing landscapes on neighborhood peers. Results of a statistical model found that change in peer water consumption resulted in a 1.7% change in individual water consumption; a 10% change in peer landscape greenness results in a 1.4% change in the household's landscape greenness. These social effects did not hold under heavily discounted irrigation water indicating that landscaping peer effects were near zero when not accompanied by an economic incentive.

13.5 Climate Change and Outdoor Water Use: What We Do Not Know?

While social scientists worked to understand the effects of lifestyle choices, public policy, and demographic trends on urban water use, natural scientists introduced climate change as an issue for urban water planning. Initially, emphasis was on the effects of a changing climate on water supplies, but more recently, concern has shifted to impacts on demand. In a study of future demand for six geographically diverse North American water utilities (Colorado Springs, Durham, Boston, Las Vegas, Tampa Bay, and San Diego), Kiefer et al. (2013) found that municipal water demand was highly sensitive to regional differences and to the variability in weather conditions. Estimated demand increased for the six cities under different climate change scenarios ranged from 1% to 12% by 2055 and from 2% to 45% by 2090, depending on the climate scenario, region, and specific utility. Climate-sensitive demands accounted for a majority of total demand in some areas and seasons. Some cities experienced increases in demand that are larger than future-projected declines in water use, suggesting that the net water use gained from conservation may not compensate for losses from climate-change impacts.

Also uncertain are the impacts of policy choices on outdoor water use. Recent efforts have emphasized drought conditions because of the obvious connection between short-term drought and long-term climate change (Karl et al. 2008). Kenney et al. (2004), for example, tracked household response to mandatory and voluntary outdoor water restrictions in six Colorado Front Range cities during drought conditions in the summer of 2002 and found that mandatory restrictions reduced water use, voluntary restrictions had limited impact, and the greatest savings occurred in cities with the most aggressive and stringent mandatory restrictions. Maggioni (2015) found that outdoor water restrictions in Southern California curbed per capita water use, but water rates and subsidies to install water-saving devices did not. In a 2007 study of severe drought conditions in six North Carolina communities, Wickman et al. (2016) found low-income households in single-family detached dwellings were more sensitive to price than high-income households were. There was, however, a more uniform response across income categories to non-price policies such as voluntary and mandatory restrictions on outdoor watering. They concluded that the burden of price increases falls more heavily on poor households. Irrigation restrictions had a more equitable impact, inducing more outdoor water conservation among high-income, high water-use irrigators.

Regulators face the need to adjust to climate change with speculative policy tools and limited capacity to anticipate potential outcomes, especially with respect to outdoor water. Beyond price, it is unclear how residents, sometimes with a deep connection to their landscaping, will respond to policy signals. As a result, water managers face deep uncertainty in planning future infrastructure and supply.

In a recent study of residential water use, De Oreo et al. (2018) concluded there is potential to reduce per capita use even further in the indoor sector. They estimated that 100% acceptance of high-efficiency devices could reduce indoor use by 35%. Additional indoor efficiency can be obtained on the customer-side through leakage control, automated metering, and leak alert programs. For a variety of reasons, as noted above, outdoor efficiencies are more difficult to anticipate as a persistent part of the population accounts for a significant proportion of wasted water. De Oreo's research involving 838 households across 22 study sites found that average outdoor use would decline by 16% if overuse was eliminated.

13.6 Problems of Deep Uncertainty

Managing the relationship between population and water (GPCD) is increasingly a problem of deep uncertainty. This relationship changed in unexpected ways over the past 65 years and is likely to change again during the next 50 years. Attempts to predict and plan water resources based on a straight-line extrapolation of the future ignored lifestyle and social issues. Added to these is the knotty problem of climatic uncertainty and the unknowns about future technology raised in the previous sections. Increasingly, water management is a problem of deep uncertainty.

Problems of deep uncertainty (often called wicked problems) involve situations in which analysts do not know or do not agree on the key drivers that will shape the future, the probability functions that describe system behavior, or how to weigh the gains and losses of particular problems. DMUU acknowledges that many aspects of the future are unknowable and that predictions and forecasts represent only one of many possible futures. DMUU does not assume that the past is an adequate guide to the future and allows for a range of potential future conditions, some of them quite dire. DMUU changes the policy question from what is the most likely future to what kind of future do we want and what policy decisions are required to achieve acceptable outcomes. This form of robust decision-making relies on community-generated scenarios of the future, focuses on the trade-offs embedded in different policy choices, and stresses community engagement with scientific modeling and decision-making processes. Emphasis is often on the search for robust solutions—those that yield an acceptable outcome across a wide range of future conditions (Lempert et al. 2003a, b).

The deep uncertainties associated with future climate impacts, water conservation strategies at state and local levels, national policy initiatives relevant to water conservation, future lifestyles, and cultural trends leave society at risk to unexpected future conditions. The Fourth National Climate Assessment (2018) articulated the central challenges to water planning and management—learning to plan for plausible futures, not only the most likely ones. Water resource management in the USA traditionally stressed a predict-and-plan approach using optimization modeling and known relationships. Some local planning agencies have adopted "what-if" approaches that explore alternative futures looking for robust planning strategies that provide acceptable outcomes over a wide range of future conditions. Las Vegas, for example used multiple demand and supply scenarios to plan for the future, engaging its citizens in public discussions of acceptable and unacceptable outcomes.

13.7 Conclusions

Many local water managers were surprised by the steady fall in water use GPCD in the recent past because they assumed a constant relationship between population and water. In fact, it took more than 10 years for the effects of the 1992 Energy Policy Act to appear in local water use records and in the USGS's Estimated Water Use series. A careful examination of this series showed three periods of GPCD: steady increases from 1950 to 1980, consistently high GPCD from 1980 to 2005, and sharp declines from 2005 to 2015. Using the most recent period to predict the future may be as problematic as assuming constant GPCD was in the recent past. Continued reductions in GPCD assume substantial reductions in outdoor use and relevant social, lifestyle, and perceptual changes associated with them. Also assumed are policy actions to change existing landscaping and an assumption that population will respond to them. Alternatively, communities can plan for an uncertain future, actively monitor ongoing population and water use, engage their populations in more active discussions of water use, and look for robust strategies that offer water security irrespective of climate change, public acceptance of conservation signals, and technological innovations. Demography is integral to water planning but not in the sense of plugging in GPCD based on past trends. Instead, it offers insights into the larger societal forces that influence not only population size and characteristics but in the way it affects water use.

References

AghaKouchak A, Feldman D, Hoerling M, Huxman T (2015) Recognize anthropogenic drought. Nature 524(7566):409–411

Arbués F, Barberán R, Villanúa I, Barberan R (2010) Household size and residential water demand: an empirical approach. J Agric Resour Econ 54:61–80

Bollinger B, Burkhardt J, Gillingham K (2018) Peer effects in water conservation: evidence from consumer migration. NBER Program (3) Environment and Energy Economics, Working Paper No 24812. https://www.nber.org/papers/w24812. Accessed 26 Feb 2019, City of Phoenix (2011)

City of Phoenix Water Services Plan. https://www.phoenix.gov/waterservicessite/Documents/wsd2011wrp.pdf

Coomes P, Rockaway T, Rivard J, Kornstein B (2010) North American residential water usage trends since 1992. Water Research Foundation. https://www.waterrf.org/PublicReportLibrary/4031.pdf

De Oreo WB, Mayer P, Dziegielwwski B, Keifer J, Sawyer PC (2018) Residential end uses of water version 2 executive report, Water Research Foundation. http://www.waterrf.org/PublicReportLibrary/4309A.pdf. Accessed 26 Feb 2019

Deiter CA, Maupin, MA, Caldwell RR, Harris MF, Ivahnenko TE, Lovelace JK, Farber NL, Linsey DS (2019) Estimated use of water in the United States in 2015. U.S. Geological survey circular 1441. https://pubs.er.usgs.gov/publication/cir1441. Accessed 26 Feb 2019

Executive Department State of California (2016) Executive order B37-16 making conservation a way of life. https://www.gov.ca.gov/docs/5.9.16_Executive_Order.pdf. Accessed 22 Sep 2017

Gober P, Quay R, Larson K (2016) Outdoor water use as an adaptation problem: insights from north American cities. Water Resour Manag 30(3):899–912

Hanak E, Mount J, Chapelle C (2016) California's latest drought. Public Policy Institute of California. http://www.adrc.asia/documents/disaster_info/2011/12/EOC_Report_16_Dec-eng.pdf. Accessed 26 Feb 2019

Höglund L (1999) Household demand for water in Sweden with implications of a potential tax on water use. Water Resour Res 35(12):3853–3863

Jackson KT (1985) Crabgrass frontier: the suburbanization of the United States. Oxford University Press, Oxford

Karl TR, Meehl GA, Miller DC, Hassol, SJ, Waple AM, Murray WL (2008) Weather and climate extremes in a changing climate. Regions of focus: North America, Hawaii, Caribbean and U.S. Pacific Islands. Synthesis and Assessment Product 3.3 Report by the U.S Climate Change Science Program and Subcommittee on Glbal Change Research. Department of Commerce and the Subcommittee on Global Climate Change Research. Department of Commerce. NOAA's National Climatic Data Center, Washington DC, USA

Kenney DS, Klein RA, Clark MP (2004) Use and effectiveness of municipal water restrictions during drought in Colorado. J Am Water Resour Assoc 40(1):77–87

Kiefer JC, Clayton JM, Dziegieleski B, Henderson J (2013) Analysis of changes in water use under regional climate change scenarios. Water Research Foundation, Denver. http://www.waterrf.org/Pages/Projects.aspx?PID=4263. Accessed 26 Feb 2019

Lempert RJ, Popper SW, Bankes SG (2003a) Shaping the next one hundred years*:* new methods *of* quantitative*,* long-term policy analysis. Rand Corporation, Santa Monica

Lempert RJ, Popper SW, Bankes SC (2003b) One hundred years long-term policy analysis. Rand Corporation, Santa Monica

Maggioni E (2015) Water demand management in times of drought: what matters for water conservation. Water Resour Res 51:125–139

Neel R, Sadalla E, Berlin A, Ledlow S, Neufeld (2014) The social symbolism of water-conserving landscapes. J Environ Psychol 40:49–56

Quay R (2015) Planning for demand: uncertainty in integrated water resource management. J Am Water Resour Assoc 107(2):32–38

Reisner M (1986) Cadillac Desert*:* the American west and its disappearing water. Viking Press, New York

Robbins P (2007) Lawn people: how grasses, weeds, and chemicals make us who we are. Temple University Press, Philadelphia

Russell S, Fielding K (2010) Water demand management research: a psychological perspective. Water Resour Res 46(50). https://agupubs.onlinelibrary.wiley.com/doi/epdf/10.1029/2009WR008408

Southern Nevada Water Authority (2018) 2018 Water resource plan & water budget. https://www.snwa.com/assets/pdf/water-resource-plan.pdf. Accessed 26 Feb 2019

United States Census (2019) Historical household tables: average population per household and family: 1940 to present. https://www.census.gov/data/tables/time-series/demo/families/households.html

Vickers A, Bracciano D (2014) Low-volume plumbing fixtures achieve water savings. Opflow 40(7):8–9

Wickman CJ, Taylor RJ, von Haefen H (2016) Conservation policies: who responds to price and who responds to prescription? J Environ Econ Manag 79L:114–134

Chapter 14
Mapping the Impact of Collaborative Research with David Plane

Beth Mitchneck

Abstract Citation counts have become a standard way of assessing the impact of a research paper. Plane and Mitchneck published two interdisciplinary and theoretically contextualized papers over 25 years ago. Plane's formal retirement marks an appropriate time to assess the impact that the papers have had in the scholarly domain. After reviewing the critical literature about citation indices and discovering that a paper with a woman, Mitchneck, as lead author has less of a likelihood of high citation than if a man, Plane, had been lead author, Mitchneck begins to count the citations of the two papers in different sources. She then conducts a qualitative assessment of the impact of the papers by reading the majority of the citations within the text in which the citation is found with the goal of understanding how the papers impacted conversations on the research topics of net migration and employment and migration during times of political and economic shock. The qualitative assessment finds contributions to conversation around the world by an international and interdisciplinary group of scholars.

Keywords Migration · Economic shocks · Citation counts · Interdisciplinary research

14.1 Introduction

Have you ever wondered what impact your research has? Have you looked at citation counts? Have you calculated your H-index? Checking out a paper's citation count or a journal's impact factor or even one's own H-index have become everyday methods of assessing impact even by avowed qualitative researchers. Many in the social, physical, natural, and life sciences and engineering use citation counts to make their cases for tenure and promotion, and, indeed, some universities require or

B. Mitchneck (✉)
University of Arizona, Tucson, AZ, USA
e-mail: bethm@email.arizona.edu

R. S. Franklin (ed.), *Population, Place, and Spatial Interaction*, New Frontiers in Regional Science: Asian Perspectives 40,
https://doi.org/10.1007/978-981-13-9231-3_14

suggest providing this information in promotion packets. While citation analysis has become a major form of research assessment at universities, we know relatively little about the dynamics and underlying influences over citation practices or how citation counts differ across measures—especially in interdisciplinary research and research conducted prior to the digital explosion of research dissemination through social media and tracking other forms of research impact. These underlying relationships undoubtedly influence any index reliant on human practices such as citations.

Now about 25 years on from the publication of two interdisciplinary papers by Mitchneck and Plane, it seems appropriate to assess their impact. What is the best way to measure the impact? As an academic, I started with citations. Little did I know that when I first started counting citations for each paper, it would lead me to seriously question the method and become even more curious about the impact of our truly collaborative and interdisciplinary research. I concluded that we needed to map out the impact rather than assess it with citation counts. The mapping included reading how other authors had cited our papers to assess whether or not we had some impact over the ways people conducted their research, did we start conversations or contribute to them, or were our papers just another one in a long list of papers on similar topics. This journey led me to conduct a qualitative analysis of the impact our papers had on related research. Below, I describe the papers that we wrote, present evidence to promote questioning the use of citation analysis alone as a measure of research impact, and then conduct a qualitative analysis of how our interdisciplinary research may have impacted other research on migration and, particularly, migration in nonmarket contexts and during economic and political shocks.

14.2 What Did We Do?

In the early 1990s, David Plane and I wrote two papers at the intersection of both of our research areas. I was an assistant professor recently hired in the Department of Geography and Regional Development (now called the School of Geography and Development). We both thought that working together would be good for my career and interesting for him to move into a new region. The time was the immediate post-Soviet period, new data on Russian migration and labor markets became available, and we could use a dataset that I had collected for a project on local economic development in Yaroslavl, Russia. Questions about the Soviet/Russian labor market that had long been considered unanswerable in Russia were now open for research. For example, was there a labor market in the new Russia, and given the nature of the emerging economy, did the labor market in Russia function as it did in a market economy? Were the demographic responses similar in the new Russia to what we experience in the United States? Although Plane and I have different specialties and approaches, we each work on topics related to human migration, and it seemed ideal to bring us together in an interdisciplinary analysis of these kinds of questions. Plane works primarily, but not exclusively, on topics related to spatial interaction and migration in the United States and from a primarily quantitative and modeling perspective. I worked on migration and economic development primarily, but not exclusively, in Russia and from a mixed methods approach both quantitative and qualitative.

14.3 The Papers

Plane and I were motivated to analyze and compare migration systems under periods of economic and political shock. While the breakup of the Soviet Union created a sensational series of shocks around the world, it was a unique period of time to establish impact on human migration that might have been useful in understanding future change. While the case is unusual, it occurred without large-scale violence (although a number of significant civil conflicts arose) and created a series of changes in the region that could trigger similar changes to migration systems.

Mitchneck and Plane (1995a), *Migration patterns during a period of political and economic shocks in the former Soviet Union*, assessed Russian migration patterns relative to past Soviet patterns and international migration systems that Soviet demographic patterns did not follow. Using newly available data on Russian migration systems, we tested a series of hypotheses that were both related to migration systems elsewhere and contextually distinct to the post-Soviet case in Russia (e.g., the impact of 15 republics becoming separate countries). We found importantly that the economic and political shocks both created new migration systems in Russia that were unique at that point in time and others that brought Russian migration more in sync with patterns that we see in market economies like the United States. We tested traditional hypotheses related to stable migration systems such as the volume, distance, rural-urban direction, and gender, age, and ethnic composition of flows. These hypotheses are linked to migration systems around the world.

Our primary finding from this study is that by using standard approaches to migration analysis with historical geographical context, we could predict the composition of flows during economic and political shock in Russia. We also highlight contextual differences, importantly the reversal of the established rural to urban labor flows due to a poorly functioning housing market and continued government attempts to manage migration by making residential permits difficult to obtain. Equally important were our findings that the shocks evident in the migration system were largely predictable and that the system was likely to continue changing due to future political and economic shifts. This paper set the scene for the use of market-based theories and analyses for this context. In addition, it also served as one of the first contextually and theoretically driven analyses of population migration patterns in the New World Order that could be reflected in other parts of the former Soviet Union and its sphere of influences, namely, Central Europe.

Mitchneck and Plane (1995b), *Migration and the quasi-labor market in Russia*, also takes a contextualized theoretical approach to investigate the relationship between migration and employment change. For example, within the Soviet Union and the post-Soviet context, do we find expected relationships between levels of net migration and employment change? Does net in-migration signify an increase in job availability? We find in this paper that employment change and net migration were not predictable in the same manner as in market economies during the Soviet period and that the economic shocks of the early 1990s made this relationship even less predictable. In a way, these findings suggest that factors other than economic ones were underlying the patterns of change that we found significant in Mitchneck and Plane (1995a). The lack of significance of employment for driving net migration in the Soviet and post-Soviet periods highlight the role that other social and political factors may have over migration systems. In particular, the social contract of full employment and the replacement of administrative determination of employment needs over the market create different incentive and information systems than one would find in a capitalist economy.

14.3.1 Summary

While the 1995 papers are distinct from one another, they share a number of qualities that make it appropriate to look at their impact together. First, both papers bring standard theoretical analysis of migration to Russia and the Soviet Union in a contextually driven way. Second, both focus on a unique period in history—the time immediately before and after the breakup of the Soviet Union. The breakup created political and economic shockwaves through the global economic and world political order as well as Russian migration systems.

14.4 Measuring the Impact of Research Papers

Measuring the impact of a research paper is important for several reasons in the academy. First, we have fallen into a pattern of checking the number of citations a paper receives so we can follow changes to our H-index (a measure of an individual's overall impact on science) or other indices which account for many of the major criticisms of the H-index—including the amount of time an individual has been publishing (see Gasparyan et al. (2018) for a strong review of the indices). Second, these indices as well as citation rates of individual papers have become standard fare in assessing the success of academics and many institutions suggest (or even require) including them in annual or promotion reviews. These indices and measures are a shorthand for assessing the impact of research because they are easy to locate and many websites calculate them for us. We are essentially using citations and the H-index as a proxy for quality rather than as a measure of quality or impact (Lehmann et al. 2006).

As I began this exercise, I collected information about citation rates for the two papers (see Table 14.1). Aware of the many critiques of citation analysis, especially around research on the social forces over how individuals cite and the lower rates of

Table 14.1 Citation counts according to major sources

Citation counts for Mitchneck Plan articles	Migration patterns during a period of political and economic shocks (1995a)	Migration and quasi labor market in Russia (1995b)
Google Scholar	33	35
ResearchGate	13	25
SCOPUS	14	20
Web of Science	13	18
Dimensions		14

citations for interdisciplinary work, I began this process with some skepticism which was only confirmed by what I found!

Table 14.1 indicates clearly that each measure of the citations for the articles shows different levels of impact. Given that the majority of research papers receive relatively few citations (see Van Noorden (2017) for an in-depth discussion of the literature on low citations and the underestimation of impact by using citation counts), I am pleased that our papers any received attention!

The literature documents both advantages and hazards of using citation analysis to assess the impact of research. The advantages are limited to easily available data and a belief that these data are comparable across individual researcher (see Gasparyan et al. (2018)). Far more research has suggested that the ways that individual researchers choose which papers and individuals to cite deems citation analysis a hazardous way to assess impact (e.g., Milard and Tanguy 2018). By hazardous, I mean that substantial research has indicated that so many different social factors impact how we choose to cite that citation indices are far from an objective source of impact and more likely a measure of fashion in science and a reflection of the demographic and disciplinary background of individual authors as well as social relations.

Let's begin with how we cite interdisciplinary research. We classify Mitchneck and Plane (1995a, b) as interdisciplinary because we combine demographic, economic, geographic, and area studies knowledge. As such, in our papers, we cite multiple specialties within our disciplines and journals outside our disciplines. This kind of interdisciplinary research generally means it is less cited and takes longer to have an impact—as measured by citations in the published literature (Van Noorden 2015).

What other factors have been documented as reducing the likelihood or at a minimum influencing article citation? Ethnicity and gender are factors in the way that we cite other research papers. Freeman and Huang (2015) in their analysis of 2.5 million authors in the United States from 1985 to 2008 found that papers with authors of different ethnicities and geographic diversity were cited more often than authors of the same ethnic background from the same geographic location. When men and women collaborate, their papers are generally cited less often than when men collaborate with other men (Beaudry and Lariviere 2016). And when the lead author is a woman, the paper is less cited than if the lead author were a man (Lariviere et al. 2013).

Research shows that the influence of gender over frequency of citation is widespread across fields. Beaudry and Lariviere (2016) find that across science and medicine, the larger the proportion of women co-authors, the lower the citations and that the same authors, when co-authoring with a male-dominated group, are cited more. A group of researchers analyzed citation patterns in international relations journals and found that women are three times more likely than men to cite work by other women and that papers written by men are highly likely not to cite work by women in international relations (Mitchell et al. 2013). Vanclay (2013) finds that in environmental science, authors can manage their citation counts by writing review articles in high impact journals.

Our professional networks also influence the ways that we cite; Milard and Tanguy (2018) find that the closer within social sphere the cited author is to the citing author, the more likely they are to cite one another. They find a significant connection between social networks and how we cite other authors. In other words, we tend to cite within our own professional networks.

The research on citations by gender and related valuation of highly cited work is not without debate. Chibnik (2014), the editor-in-chief of the *American Anthropologist* at that time, finds that there are no statistically significant citation patterns by gender in the journal. Others note that self-citation by prolific authors contributes to skewing the data and that men tend to self-cite more than women (Cameron et al. 2014). The fact that substantial attention is now being paid to developing alternative ways to assess the impact of one's research suggests that there is broader consensus around the scientific conclusion that more citations means more impact.

The perspective that assessing article value should focus on alternative measures through how that research is disseminated (Fenner and Lin 2014) is gaining popularity. There is recognition that a New World Order has developed around measuring impact that ties to control for the type of attention an article receives, through what mechanisms and then what impact really means. Altmetrics (https://www.altmetric.com) is an excellent means to map out the impact of one's work in a twenty-first-century world including social media and public policy reports. While opening up the assessment of research impact to a larger variety of sources other than journal citations is a step in the right direction, assessment of the impact of social factors on how individuals choose what to cite should continue.

Considering the widespread influence of factors unrelated to the quality of research over citation rates, Mitchneck and Plane (1995a, b) should not receive many citations at all! We published before the Internet explosion of social media and altmetrics, a woman is the lead author, we come from the same racial and geographic profiles, and the work is interdisciplinary. Given how the cards are stacked against us, I am thrilled that our work is cited at all!

14.5 Mapping Research Impact: A Qualitative Assessment

Using traditional citation analysis is not likely to provide a deep understanding of the impact of our work in the 1990s. Knowing this, and still wanting to assess the impact of our research, I conducted a qualitative analysis of how the two papers were cited rather than the number of times. Using the literature on citation analysis, I developed a methodology to map out the impact of the research according to a number of criteria that would be meaningful to us and that we could assess. To assess the audience we reached, I collected information based on what we know about the influences over who we cite in our research. Were we reaching the same people we are always in a conversation with? If so, we would know the majority of those who cited our work. Were we reaching the targeted interdisciplinary and international audience, one that was interested in similar topics like what happens to people in an emerging economy or under political and economic shocks? If the work was truly interdisciplinary that too would be reflected in the characteristics of who cited us and how. Assessing how we were cited would get at the issue of how our work impacted other research on the topic.

I used both ResearchGate and Google Scholar to document the number of citations and detailed information on the author and the publication in which our papers were cited. Using the literature on citation analysis and the underlying forces that influence how we cite, I collected data on the title, author(s), year of publication, journal or book name, discipline of the author, topic and region of the paper, and whether or not we knew the author(s). I then read the portion of the text where our research was cited (when possible through the internet) to assess by whom and how we were cited. To see if the citations signified some positive quality of our work, I assessed whether or not our work was cited to simply document that work is done in our subject area, or if we were part of a conversation such as to support a finding, tell a story, or summarize our research. The latter would suggest that our work had some impact on research that came after our papers. I also assessed if the authors are known to us or if we self-cited (the literature suggests that both would increase our citation count). Finally, wanting to assess whether or not our research was viewed as off base, I read to make sure that we were not at the center of any debate or nasty disagreement.

14.6 Findings

By whom were we cited? My analysis suggests that we were cited by a diverse group of authors who create a map that is highly international in terms of where the authors live and work and the regional topics on which they published. Our map of influence extends really around the Western world from the United States (e.g., Heleniak 2009) and Canada (e.g, Lo and Teixeira 1998) throughout Europe (Russia, Greene (2012); Germany, Lerch (2014); Norway, Gentile (2006); Finland, Lonkila and Salmi (2005)). The fields that the citing authors come from include demography, economics, geography from various subdisciplines, history, political science, sociology, and urban planning.

How were we cited? My analysis indicates that we were cited in ways to suggest that we impacted conversations about the effects of economic and political shocks on migration systems and how information about employment influences migration systems. Those were our intentions! As expected, we were cited in a variety of ways including being part of a list of citations on the topic of internal migration (e.g., Gerber 2006; Lonkila and Salmi 2005) and economic disruption (Curran et al. 2016).

More importantly, we were cited in diverse ways such as in a paper about the possible breakup of Canada (Lo and Teixeira 1998) in terms of both our methodology and findings. In a note, Lo and Teixeira (1998: 495) write:

> In an empirical analysis of migration in Yaroslavl Oblast, a region in central Russia, Mitchneck and Plane (1995a) examine how severe economic and political shocks due to disintegration of the former Soviet Union might change the Russian migratory system. While the data clearly show a migration system undergoing shocks, the authors found the structure of the system still predictable using analyses of historical trends and standard approaches.

Another paper about Albania (Lerch 2014: 1535) cited our research to support structural reasons for migration in a transition context:

> Structural effects can be expected to a lesser extent, particularly in the first decade of transition in the societal system, which motivated undifferentiated migration in other post-communist contexts. (Mitchneck and Plane 1995a)

Our paper on employment and net migration (Mitchneck and Plane 1995b) was also cited in substantive ways. For example, Fan (2005: 296) notes:

> Even after the late 1980s, mobility in Russia was still unduly affected by the legacy of the Soviet-period registration system and access to services and resources tied to that system (Mitchneck and Plane 1995b). Likewise, migration control exists in China (covered in the next section).

And Greene (2012: 138) writes citing to Mitchneck and Plane (1995b):

> Beth Mitchneck and others remind us of the importance of one's workplace for the provision of social services and, indeed, for the maintenance of one's entire lifestyle during the Soviet period, and argued that the continual provision of such services through the workplace in the early transition period acted as a brake on labor migration.

In what geographic context was our work cited? Our work was cited in publications about migration in Albania (Lerch 2014), Estonia (Sjoberg and Tammaru 1999), Russia and Central Asia (Sahadeo 2013; Earle and Sabirianova (2002), and others), Thailand (Curran et al. 2016), Asia and the Pacific (Skeldon 1998), Hungary (Zueva 2005), China (Fan 2004 and others), and the Soviet Union (Eastman 2013). The map of our papers' influence extends beyond the borders of the authors to places in Asia that are experiencing similar processes to what happened in Russia.

Our papers were cited in mainly interdisciplinary journals supporting our intention to target an interdisciplinary audience. As noted above, the authors come from many different disciplinary backgrounds. Also supporting the interdisciplinary nature of the work is the slow to gain citations noted in *Nature* about interdisciplinary research (Van Noorden 2015). Mitchneck and Plane (1995a) has roughly half of its citations after the 13-year mark, and Mitchneck and Plane (1995b) has about only one-third. Google Scholar did not have complete citation information for all of the citations. Mitchneck and Plane (1995b) about the quasi-labor market continues to be cited, most likely because scholars continue to have interest in the functioning of the Soviet labor market.

Self-citation does not seem to play a role in our citation counts. Our own citations to this work account for only four of the over sixty citations.

14.7 Summary and Implications

Citation analysis provides a relatively simple method of assessing the impact of a scholarly paper. The literature shows, however, that by only looking at citations in other scholarly publications, we undervalue the impact of a paper. As Lehmann et al. (2006: 1004) so eloquently state, "Unfortunately, the potential benefits of careful citation analyses are overshadowed by their harmful misuse." Altmetrics are an improvement on how we measure impact because they include a variety of different channels for assessing impact beyond journal or book publication. Yet, there is substantial evidence that points toward how social forces mold the ways that the original citations and sharing of our research occurs related to gender and ethnicity.

In the case of our two papers, the fact that the lead author is a woman and that a man and a woman from the same general discipline or interdiscipline are co-authors suggests that our papers will be undercited or at a minimum cited less than if Plane were the lead and if he had written the paper with another man! Being interdisciplinary further disadvantages the likelihood for citation. Yet, the qualitative analysis of who and how our papers are cited show that the papers had a wide reach in terms of geography and discipline and that they continue to be cited 26 years after publication. The textual read of how we were cited shows that most often our papers were cited in a substantive way rather than documenting that it was written on a topic about which other authors are also writing. In that sense, these papers have had impacts on conversations in research about Russia and the Soviet Union and contributed to conversations about many other countries that experience political and economic shock.

What are the implications of this analysis of the impact of the two papers? Clearly, citation counts are not nearly enough information to assess the impact of research or the productivity of a scholar. The H-index has had enough criticism in the literature. But citation counts have not. Can everyone read through and assess the ways in which their papers have been cited to document the impact of our work? Probably not, but it does document some important areas of assessment in the academy such as high values given to the map of our research—the geographic extent of our reach/reputations and interdisciplinarity. Yet, if we are going to use counts and indices toward faculty advancement practices, we need to better understand the advantages and hazards of citation counts. Perhaps an in-depth analysis such as I have done here would be useful for one or two papers in a promotion package. The use of alternative metrics for newer publications is an excellent way to include a larger variety of ways that our research can impact conversations and public engagement.

In his career, Plane has contributed to many conversations and has had a global reach through his research and his leadership in scholarly associations. The mapping of these papers fills out that map in new ways.

References

Beaudry C, Larivière V (2016) Which gender gap? Factors affecting researchers' scientific impact in science and medicine. Res Policy 45(9):1790–1817

Cameron EZ, Edwards A, White AM (2014) AM Halt self-citation in impact measures. Nature 505:160

Chibnik M (2014) Gender and citations in American anthropologist 116:493–496

Curran SR, Meijer-Irons J, Garip F (2016) Economic shock and migration: differential economics effects, migrant responses, and migrant cumulative causation in Thailand. Soc Dev 2 (2):119–157

Earle JS, Sabirianova KZ (2002) How late to pay? Understanding wage arrears in Russia. J Labor Econ 20(3):661–707

Eastman JT (2013) Youth migration, stratification and state policy in post-Soviet Russia. Sociol Compass 7(4):294–302

Fan CC (2004) The state, the migrant labor regime, and maiden workers in China. Polit Geogr 23 (3):283–305

Fan CC (2005) Interprovincial migration, population redistribution, and regional development in China: 1990 and 2000 census comparisons. Prof Geogr 57:295–311

Fenner M, Lin J (2014) Novel research impact indicators. LIBER Q 23(4):300–309

Freeman R, Huang W (2015) Collaborating with people like me: ethnic coauthorship within the United States. J Labor Econ 33(s1):s289–s318

Gasparyan Y, Yessirkepov M, Duisenova A, Trukhachev VI, Kostyukova EI, Kitas GD (2018) Researcher and author impact metrics: variety, value, and context. J Korean Med Sci 33(18): e139. https://doi.org/10.3346/jkms.2018.33. e139 accessed online March 20, 2019

Gentile M (2006) From migration to segregation in the former closed city. Geogr Pol 79(2):23

Gerber TP (2006) Regional economic performance and net migration rates in Russia, 1993–2002. Int Migr Rev 40(3):661–697

Greene SA (2012) Citizenship and the social contract in post-Soviet Russia: twenty years late? Demokratizatsiya 20(2):133–141

Heleniak TE (2009) The changing spatial distribution of the population of the former Soviet Union (Doctoral dissertation)

Altmetrics https://www.altmetric.com Accessed March 19, 2019

Larivière V, Chaoqun NI, Gingras Y, Cronin B, Cassidy SR (2013) Global gender disparities in science. Nature 504:211–213

Lehmann S, Jackson AD, Lautrup BE (2006) Measures for measures. Nature 444:1003–1004

Lerch M (2014) The role of migration in the urban transition: a demonstration from Albania. Demography 51(4):1527–1550

Lo L, Teixeira C (1998) If Quebec Goes . . . The 'Exodus' impact? Prof Geogr 50:481–498

Lonkila M, Salmi AM (2005) The Russian work collective and migration. Eur Asia Stud 57 (5):681–703

Milard B, Tanguy L (2018) Citations in scientific texts: do social relations matter? J Assoc Info Sci Tech 69(11):1380–1395

Mitchell SM, Lange S, Brus H (2013) Gendered citation patterns in international relations

Mitchneck B, Plane DA (1995a) Migration and the quasi-labor market in Russia. Int Reg Sci Rev 18 (3):267–288

Mitchneck B, Plane DA (1995b) Migration patterns during a period of political and economic shocks in the former Soviet Union: a case study of Yaroslavl' oblast. Prof Geogr 47(1):17–30

Sahadeo J (2013) The accidental traders: marginalization and opportunity from the southern republics to late Soviet Moscow in movement, power and place in Central Asia and beyond. Routledge, pp 153–172

Sjoberg O, Tammaru T (1999) Transitional statistics: internal migration and urban growth in post-Soviet Estonia. Eur Asia Stud 51(5):821–842

Skeldon R (1998) Urbanization and migration in the ESCAP region. Asia-Pac Popul J 13(1):3

Van Noorden R (2015) Interdisciplinary research by the numbers, an analysis reveals the extent and impact of research that bridges disciplines. Nature 525:306–307

Van Noorden R (2017) The science that's never been cited. Nature 5552:162–164

Vanclay J (2013) Factors affecting citation rates in environmental science. J Informet 7(2):265–271

Zueva A (2005) Gendered experiences in migration from Russia to Hungary in the 1990s multicultural center Prague. www.migrationonline.cz