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