geographical analysis

Geographical Analysis ISSN 0016-7363

Determinants of Variation in Population–Employment Interaction Findings: A Quasi-Experimental Meta-Analysis

Gerke J. Hoogstra,[1] Jouke van Dijk,[1] Raymond J. G. M. Florax[2,3]

[1]Urban and Regional Studies Institute (URSI), Faculty of Spatial Sciences, University of Groningen, Groningen, The Netherlands, [2]Department of Agricultural Economics, Purdue University, West Lafayette, IN, [3]Department of Spatial Economics, VU University–Amsterdam, Amsterdam, The Netherlands

This article reports about a metaregression analysis of empirical results generated using data for the northern Netherlands (1988–2002) in order to investigate the ambiguity in results in the population–employment interaction literature. Specifically, the analysis deals with the issue whether ''jobs follow people'' or ''people follow jobs.'' The article starts with introducing the basics of quasi-experimental meta-analysis and with identifying some advantages of using quasi-experimental meta-analysis as compared with the standard meta-analysis approach. Two subsequent sections document the selection of the population–employment interaction model and salient characteristics of the data set as well as the setup of the primary analyses. A total of 4,050 quasi-experimental empirical results for the jobs–people direction of causality are generated using different specifications and estimators for a spatial econometric interaction model. The subsequent metaregression analysis reveals that the empirical results are largely shaped by the spatial, temporal, and employment characteristics of the data sampling. The results also appear much more sensitive to different measurements of the model's key variables when compared with alternative specifications of the spatial weights matrix. The main determinant driving empirical results about jobs–people causality are differences in model specification and estimation, as revealed by an inherent bias in parameter estimates and misguided inferences for some of the commonly used specifications. Finally, suggestions for future research are identified.

Introduction

The question whether ''jobs follow people'' or ''people follow jobs'' has generated a vivacious discussion in the population–employment interaction literature. An obvious circumstance that triggered the debate is the allegedly striking discrepancy

Correspondence: Jouke van Dijk, Urban and Regional Studies Institute (URSI), Faculty of Spatial Sciences, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands
e-mail: jouke.van.dijk@rug.nl

Submitted: December 1, 2005. Revised version accepted: July 30, 2009.

in the results of different empirical studies of the topic, which is well documented in both qualitative and quantitative reviews of the literature (Bollinger and Ihlanfeldt 2001; Hoogstra, Florax, and van Dijk 2005). Naturally, questions have been raised about why research findings about the jobs–people direction of causality appear so overly varied. One reason could be that the findings epitomize real-world differences in this empirical phenomenon. In this case, applications for different areas and time periods inevitably lead to divergent research findings beyond what should be expected on the basis of random sampling variation per se. Alternatively, as pointed out by Boarnet, Chalermpong, and Geho (2005) and others, the mixed empirical evidence may not signal real-world variation; rather, it may represent a scientific artifact stemming from methodological differences between studies.

In the literature about population–employment interaction, virtually all studies provide a host of estimation results that can furnish preliminary insights into the effects of using a particular type of data or methodology. Given the large array of variations and the complex nature of potential interactions across variations, the identification of the robustness of study results against variations in an underlying study's characteristics requires a more rigorous assessment. Meta-analysis, constituting a set of statistical tools to synthesize research results and to identify important features explaining the variation across research results, is particularly suited for such a robustness analysis (Stanley and Jarrell 1989; Stanley 2001). Although meta-analysis is typically used to analyze what constitutes the state-of-the-art or the bottom line of an existing body of studies, meta-analytical techniques also can be used to give a systematic statistical account of research findings obtained in a quasi-experimental setup (Florax and de Graaff 2004). In general, meta-analysis is used to analyze a sample of study results obtained for different data sets. The quasi-experimental approach applied in this article involves repeated sampling from a single, well-known data set to generate artificially study results by systematically varying the specification and other aspects of the research setup. In this case, meta-analytical techniques are utilized to provide a robustness or sensitivity analysis. The existing empirical literature is used merely to identify useful dimensions of variation in research setup that should be included in the experiments (Florax, de Groot, and Heijungs 2002; Banzhaf and Smith 2007).

In this article, the quasi-experimental meta-analysis technique is used to test differing hypotheses about the empirical nature of population–employment interaction as well as to identify to what extent the methodological setup of underlying (quasi) studies impacts research findings. Specifically, the robustness of findings about the jobs–people direction of causality is assessed for three substantive and three methodological study features that have been intensively debated in the literature. The substantive study features are concerned with the extraction of different samples from a data set that contains detailed temporal, spatial, and sectoral employment information. The analysis provides substantive insights into whether empirical results about the jobs–people direction of causality differ over time, over space, and between employment groups.

Methodological study features relate to the operational definition of variables, the specification of the spatial weights matrix, and model specification and estimation to identify whether these features play a role in shaping reported research findings. The selected features are particularly relevant because the operational definition of variables and the specification of the weights matrix show considerable variation across studies in this literature (Hoogstra 2011). Moreover, the possible impact of these features already has been the subject of some preliminary inquiries (notably by Mulligan, Vias, and Glavac 1999; Henry, Schmitt, and Piguet 2001; Boarnet, Chalermpong, and Geho 2005). Findings from these studies clearly highlight the need for further investigation. With regard to model specification and estimation, the impact of accounting for spatial dependence in the form of a spatially lagged dependent variable is investigated. Although erroneous omission of these spatially lagged variables causes omitted variable bias, this specification issue has been ignored for a very long time because of complications in the estimation of such models. However, the availability of a feasible generalized spatial two-stage least-squares estimator (GS2SLS), which is straightforward in its implementation (Kelejian and Prucha 2004), facilitates identifying the impact of differences in the specification and estimation of alternative spatial models.

The impact of the selected study features is determined by repeated testing for the nature of population–employment interaction, yielding as many as 4,050 quasi-experimental research findings. These findings are generated by means of systematic variations in a spatial econometric interaction model estimated for a single database of observations from the northern Netherlands (1998–2002), using all possible combinations of a particular data selection (by time period, region type, and employment type), variable measurement, spatial weights matrix design, and model specification and estimation.

Quasi-experimental meta-analysis

The quasi-experimental approach adopted for this study resembles a standard meta-analysis in the sense that statistical techniques are applied to assess the robustness of a collection of study results against a variety of study characteristics. The main distinction is that the metadata are taken not from a series of primary studies using different data sets but from an exhaustive series of quasi-study results generated using one specific data set. The quasi-experimental approach circumvents some of the pitfalls associated with a standard meta-analysis (Florax, de Groot, and Heijungs 2002). For instance, in the case where a meta-analysis is built on aggregate statistical summary indicators from a compilation of studies, difficulties usually arise from the heterogeneity of the underlying studies. Furthermore, uncovering features responsible for the variation in research outcomes often turns out to be rather difficult, because of strong correlations between the underlying study characteristics. Finally, closely related to the previous arguments, the novelties of individual studies are not always reproduced in great numbers later on. With replications largely lacking, the existing literature may simply not exhibit sufficient variation to permit a meaningful statistical analysis that can identify the impact of study characteristics responsible for the variation in research outcomes.

By applying meta-analytical techniques to research findings that are obtained in a quasi-experimental setup, the preceding described difficulties are partly mitigated. Instead of being at the mercy of the limitations and possibly limited availability of existing studies, the literature is used to identify the variations in study features that should be investigated in a quasi-experimental setup. By having complete control over the data-generating process in the latter setup, unobserved heterogeneity across studies should not be a problem as long as features that are not considered central to the analysis are kept constant across experiments. Similarly, potential problems due to multicollinearity or lack of variation are easily evaded, providing all, or at least a large number, of the possible combinations of the principal study features are utilized in the series of experiments. By allowing for direct control over the setup of the experiments, the quasi-experimental approach can be tailored to statistical inference in the meta-analysis to allow for proper identification of the impact of relevant study features.

The quasi-experimental approach is akin to response surface techniques developed in econometrics. Basically, these techniques hinge on the estimation of an auxiliary regression in which each observation corresponds to one experiment. The dependent variable reflects some estimated output quantity of the experiments, whereas the independent variables reflect the research dimensions that have been allowed to change across experiments. Strictly speaking, each experiment extends only to the data-generating process that underlies that particular experiment, and a series of experiments extends merely to a finite set of data-generating processes. However, by combining the various experiments in a response surface, the results can be generalized to a larger population of data-generating processes (Davidson and MacKinnon 1993; see also Florax and de Graaff 2004).

Quasi-experimental meta-analysis (or response surface analysis) has proved to be a valuable tool to evaluate the sensitivity of research outcomes to alternations in research setups. For example, it has been used with some success in the economic literature about gross domestic product growth to settle the seemingly endless list of growth determinants (see, e.g., Florax, de Groot, and Heijungs 2002 and the references therein). Besides the examination of substantive issues, such as the sources of economic growth, it has become commonplace in methodological studies that need to summarize the abundant output of Monte Carlo experiments (see, e.g., Dubin 2003). Banzhaf and Smith (2007) observe that the potential of meta-analysis is not limited to such designed experiments but instead stretches to practically any research in which modeling judgments are made. Although understanding the robustness of findings is clearly important for assessing empirical work, and in many cases may be a separate source of insight, space limitations together with a desire to avoid the appearance of ''data mining'' mean that the role of such judgments is rarely documented. Meta-analysis, Banzhaf and Smith (2007) argue, allows researchers to document concisely and to explicate the impact of the judgments underlying their research, so that fellow practitioners also can benefit from the insights gained from model development that would otherwise have remained sorrowfully hidden.

Research design: econometric model and data

The general framework to be used in the quasi-experimental primary studies generated for this meta-analysis is Boarnet's (1992) spatial econometric version of the classic simultaneous equations system with adjustment lags introduced by Carlino and Mills (1987). Over the years, the Carlino–Mills (CM) model has been the standard for investigating population–employment interaction and has been adopted in over 50 studies, most geared toward the United States (see Hoogstra 2011 for an overview). In broad terms, a distinction can be made between interregional studies that have focused on counties (or county aggregates) as the spatial units of observation (see, for instance, Carlino and Mills 1987; Mulligan, Vias, and Glavac 1999; Carruthers and Vias 2005; Vias and Carruthers 2005; Carruthers and Mulligan 2007, 2008) and intraregional studies that have examined the distribution of jobs and people at a finer spatial scale, such as at the municipality or census tract level (see, for instance, Boarnet 1992, 1994; Bollinger and Ihlanfeldt 1997; Henry, Schmitt, and Piguet 2001; Boarnet, Chalermpong, and Geho 2005; Schmitt et al. 2006). The Boarnet model differs from the regular CM model in that the spatial units under examination are no longer assumed to match regional labor markets in which population–employment interaction operates. Instead, it adjusts for the possible spatial mismatch between these units and actual labor market zones by allowing the interplay between population and employment to stretch beyond the boundaries of single observation units.

Thus, while the observational units of an intraregional analysis usually are less similar to actual labor market zones than the corresponding units of an interregional analysis, and therefore, probably less suited for investigating the issue of population–employment interaction per se, intraregional analyses allow explorations of the salient concerns about spillover effects among these units, the technical issue of how to control for these effects, and ultimately whether the spatial econometric technique being adopted affects the results in a substantial manner. Such understanding is especially useful for small area models that aim to understand local development patterns, as in large parts of the urban economics literature. The Boarnet model, which, because of the peculiarities associated with small area observations, provides the most interesting case for further exploration, is formally given by the following equations:

$$\Delta P_{i,t} = \alpha_0 + \alpha_1 X_{i,t-1} + \alpha_2 P_{i,t-1} + \alpha_3 EMP_{i,t-1} + \alpha_4 \Delta EMP_{i,t} + u_{i,t}, \tag{1a}$$

$$\Delta E_{i,t} = \beta_0 + \beta_1 Y_{i,t-1} + \beta_2 E_{i,t-1} + \beta_3 POP_{i,t-1} + \beta_4 \Delta POP_{i,t} + v_{i,t}. \tag{1b}$$

where $P_{i,\,t-1}$ is the population size in location i at year $t-1$, $\Delta P_{i,\,t}$ is the population change in location i between t and $t-1$ as defined by $P_{i,\,t} - P_{i,\,t-1}$, $POP_{i,\,t-1}$ is the population size of i's labor market zone, $\Delta POP_{i,t}$ is the population change in location i's labor market zone, and $X_{i,t-1}$ ($Y_{i,t-1}$) is a vector of population– (employment) related location characteristics of i, preferably measured at time $t-1$ to avoid simultaneity bias. A similar set of definitions holds for the employment indicators E and EMP. Additionally, α_k and β_k are parameters to be estimated, and $u_{i,t}$ and $v_{i,t}$ denote stochastic errors.

The pivotal feature of the spatial econometric model proposed by Boarnet is the inclusion of the right-hand-side labor market variables, which are obtained by means of a spatial lag operation. This operation involves recomputing the population and employment values of individual locations in conjunction with those of their neighbors, as specified by a spatial weights matrix W. For a set of n observations, the matrix W is an $n \times n$ positive matrix in which $w_{ij} \neq 0$ defines j as being a neighbor of i, and $w_{ij} = 0$ otherwise. By convention, the elements of the diagonal are set to zero. The weights structure implied by the specification of matrix W rests on contestable assumptions about the spatial arrangement of the data at hand and can take on a variety of forms. In the meta-analysis, three alternative weighting schemes are used. Formally, the labor market variables are given by $POP = (I+W)P$, $EMP = (I+W)E$, $\Delta POP = (I+W)\Delta P$, and $\Delta EMP = (I+W)\Delta E$, where I is the $n \times n$ identity matrix. Premultiplying by I adds then location values, which, due to the zeros on the main diagonal of W, have been excluded otherwise. In case the row elements of W are standardized, such that they add up to one (which is commonly preferred as it facilitates interpretation and comparison [Anselin 2002]), the labor market variables measure the sum of a location's population or employment values and (weighted) averages of the corresponding values in neighboring locations.

Of particular interest for this study are the endogenous labor market variables ΔPOP and ΔEMP, the parameters of which reveal the nature of population–employment interaction.[1] A statistically significant positive estimate for α_4 points to "people follow jobs," whereas a statistically significant positive estimate for β_4 points to "jobs follow people." Dual causality or two-way interaction is confirmed when both parameters reveal the same, positive, sign and they are statistically significantly different from zero. The particular form of spatial simultaneity introduced by the spatial lag of the dependent variable of each equation appearing on the right-hand side of the other equation has been termed a spatial cross-regressive model (Rey and Boarnet 2004) and implies that the estimation of α_4 and β_4 is not without complications. For instance, obtaining the predicted rather than the observed values for ΔPOP and ΔEMP in the first stage of a routine two-stage least-squares (2SLS) estimation procedure requires different procedures. In this study, we adopt a technique previously used by Bollinger and Ihlanfeldt (1997) and Henry, Schmitt, and Piguet (2001), in which these values are directly obtained by using all of the model's predetermined variables, plus their spatial lags and higher-order spatial lags up to the order of three (hence, W, W^2, and W^3). According to Rey and Boarnet (2004), the chosen technique compares favorably with the traditional method of obtaining predicted values for ΔE and ΔP (by using the predetermined variables but without their spatial lags as instruments), which then are multiplied by matrix W to obtain the predicted values for ΔEMP and ΔPOP. The latter technique yields biased estimates in the likely event that these instrumented spatially weighted variables correlate with the residuals. In contrast, the approach adopted here ensures by construction that these variables are orthogonal to the residuals.[2]

A final key issue is whether other forms of spatial dependence are present in the system of equations that must be accounted for. For instance, the data-generating process may be such that spatial dependence also exists in the dependent variables (in addition to the right-hand-side endogenous variables), a complication that can be remedied by including the spatial lag of these variables on the right-hand side of the equation, or what has been called a spatial autoregressive (SAR) term. The cross-regressive model described by equations (1a) and (1b) is just one of several alternative spatial econometric models for simultaneous equations systems (see Rey and Boarnet 2004 for an overview and Henry, Schmitt, and Piguet 2001 for applications), which also include an ''augmented Boarnet model'' in which these autoregressive terms are added. Formally, this model (called B-SAR by Henry, Schmitt, and Piguet 2001) is given by the following equations:

$$\Delta P_{i,t} = \rho W \Delta P_{i,t} + \alpha_0 + \alpha_1 X_{i,t-1} + \alpha_2 P_{i,t-1} + \alpha_3 EMP_{i,t-1} + \alpha_4 \Delta EMP_{i,t} + u_{i,t}, \quad (2a)$$

$$\Delta E_{i,t} = \gamma W \Delta E_{i,t} + \beta_0 + \beta_1 Y_{i,t-1} + \beta_2 E_{i,t-1} + \beta_3 POP_{i,t-1} + \beta_4 \Delta POP_{i,t} + v_{i,t}. \quad (2b)$$

Crucially, in the presence of spatial dependence in the dependent variables, least-squares estimations of the model described by equations (1a) and (1b) yield biased and inconsistent parameter values, including those for α_4 and β_4. But while the B-SAR model described by equations (2a) and (2b), in contrast, does allow for spatial dependence in the dependent variables (and thus is preferred), it has long been hindered by the lack of an appropriate estimation technique. However, Kelejian and Prucha (2004) recently suggested a GS2SLS estimation procedure that has proved to yield consistent and asymptotically normal parameter estimates for this preceding case, in which spatial dependence exists in both the dependent variables and the right-hand-side endogenous variables, or put simply, simultaneity in the presence of spatial dependence. Here, together with the issue of model specification (excluding versus including the SAR lag), the impact of using this technique on the parameter estimates that indicate whether jobs follow people or people follow jobs is systematically compared with that of routine simultaneous equations estimations.

For the estimation of the model parameters, and α_4 and β_4 in particular, we use a cross-section sample of settlement-level data from Fryslân, a province in the northern part of the Netherlands (see Fig. 1 for a map of the study area). Only 1,275 mi^2 in size, the study area contains no less than 392 settlements, each including an inner built-up area that is primarily surrounded by agricultural land. On average, these settlements are about 3.2 mi^2, which is about the same size as U.S.

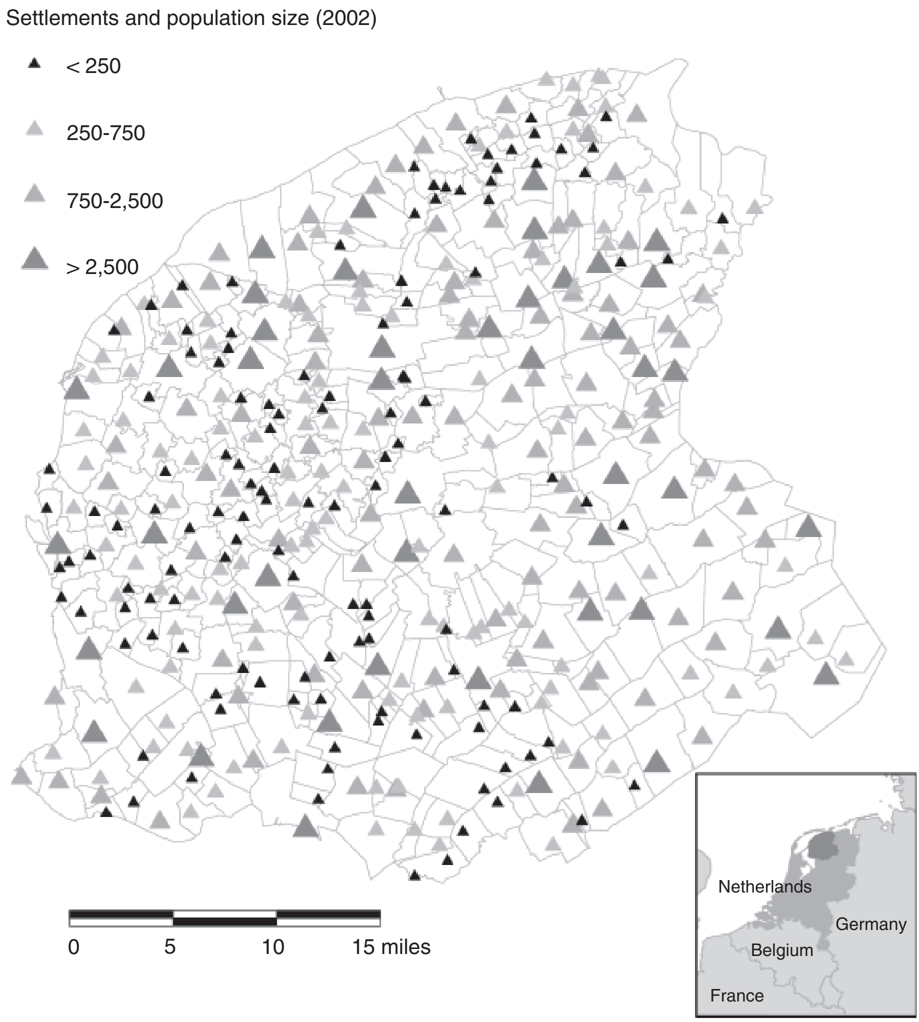

Figure 1. Study region (Province of Fryslân, the Netherlands) including 392 settlement areas.

census tracts or French ''communes'' examined in previous population–employment interaction studies (see Henry, Schmitt, and Piguet 2001; Boarnet, Chalermpong, and Geho 2005; Schmitt et al. 2006). A unique commuting flows data set for this region (Van der Horn, Hellingwerf, and Veldhuis 2001) reveals that 54% of these flows are across different settlements, which makes these units ideal for investigation in a spatial econometric model. Also, the conditions for working and living in this region are highly fragmented, meaning that households and firms have a great deal of choice in evaluating locations, even when a decision is reduced to a relatively small geographical area within the region. Finally, the study region is economically very much internally oriented, being dominated by small- and medium-sized firms that mainly serve local and regional markets, with interindustry linkages being maintained by local firms in particular (see RUG/CBS 1999) and with local residents holding more than 95% of all full-time jobs. Hence, this study region appears to be well suited for investigation on its own, which obviously eases the identification of the factors determining spatial change.

The main data used in this study, relating to the period 1988–2002, measure the employment and population size of settlements as the total number of full-time jobs in local establishments (which include all activities, except agriculture) and local residents, respectively. Besides these essential population and employment data, data are needed for the exogenous variables (X and Y), which for convenience are kept the same across the different model estimations. Hence, data are selected that capture some of the salient settlement-specific characteristics and that one can reasonably assume have not significantly changed over the 14-year period under examination. In short, we include two variables that describe the age structure of the settlements' residents (i.e., the proportion of people younger than 15 and older than 64 years, respectively), and one composite variable based on income, education, and unemployment, which can be seen as a proxy for social status (see Knol 1998). Two variables capture the access to important transport junctions as measured by the straight-line distance to the nearest railway station and motorway entrance/exit point, whereas data about the actual travel time by car to the capital city of the Netherlands (Amsterdam) are used to measure the relative location of these settlements in a national context. Next, a dummy variable is included to proxy the "Regional Plan" (*Streekplan*) of 1994, which has been the principal instrument used by provincial authorities to impose their policies about land use and the location of jobs and people. Specifically, the dummy is set to one for settlements being located in specially designated economic growth zones and individual settlements with important recreational functions, and it is set to zero for the remaining settlements. Next, we include a dummy variable to control for possible spillover effects from outside the research area. Specifically, the dummy is set to one for the border settlements where a significant proportion of the local residents works in neighboring provinces and set to zero for the remaining interior settlements. Finally, following suggestions by Boarnet, Chalermpong, and Geho (2005), we use detailed information about land use patterns to improve the reliability of the lagged adjustment parameters (an issue that is not further addressed in this study; see note 1). Specifically, six variables are included that measure the area in each settlement for the respective land use categories: agriculture, forests, nature, water, recreation, and infrastructure.

Experimentations

Data selection issues

Among the study features to be examined in the meta-analysis, we select three substantive ones that reflect the inner workings of population–employment interaction. Specifically, questions about the *empirics* of the jobs–people direction of causality are addressed by using different subsamples of the data. First, the model described in ''Research design: econometric model and data'' is repeatedly tested for different time periods to uncover whether the jobs–people direction of causality is subject to temporal changes. A priori, the idea of time effects seems very intuitive because the preferences for business and residential location, and economic conditions to act upon these preferences, change over time. Such effects may reflect (short-term) business cycle fluctuations as well as fundamental (long-term) societal changes. For example, the common assumption about the transformation from an industrial-based society to a knowledge-based society is that the balance is increasingly shifting toward jobs following people instead of the other way around (see Florida 2002). The assumption of temporal shifts already has encouraged researchers to generate multiple estimation results for different time periods. In the most detailed study, Mulligan, Vias, and Glavac (1999) report considerable variation in the results on population–employment interaction across one-year time periods between 1969 and 1994. Meanwhile, the combination and comparison of results *across* studies for different time periods, as assessed through a routine meta-analysis of CM studies (see Hoogstra, Florax, and van Dijk 2005), does not reveal any clear-cut time effect, although the lack of sufficient variation among these studies may be partly responsible for this result. Here we divide the population and employment data into six different, partly overlapping, four-year time periods between 1988 and 2002 to make a more accurate assessment of possible temporal shifts in the direction of causality than can be done by a standard meta-analysis. We selected a four-year time lag because the model is essentially a long-term growth model for which the use of, for instance, a one-year time lag is not particularly suited. The focus on a long time span, such as the 10-year lag often used in the literature, appears to mask too much of the varying circumstances during the 1988–2002 period, which includes a turning point in the business cycle around 1994. Somewhat in between, the use of a four-year time lag suggests a useful comparison of population–employment interaction over time and, besides, ensures that a significant number of observations can be looked upon in the subsequent meta-analysis.

The second aspect of data sampling to be investigated concerns the spatial nature of population–employment interaction. Widespread support exists among researchers for the thus far largely untested claim that estimation results supporting the jobs–people direction of causality exhibit spatial nonstationarity. Being one of the most eye-catching differences between studies, the ''region under examination'' is typically viewed as one of the main sources for the substantial variation in research findings that typifies this literature. However, fundamental reasons also exist to suggest that the findings are shaped by the geographical characteristics of data, which basically implies that conditions for working and living are not the same everywhere. Findings by Schmitt et al. (2006) indicate that even within a seemingly coherent group of French rural areas, considerable differences can be found in the direction of population–employment interaction, depending on factors such as the size and growth of a nearest urban center. Likewise, separate model estimations by Boarnet (1992), with a complete data set of New Jersey municipalities and a subset of slow-growing municipalities, hint at the presence of what he refers to as ''structural breaks in the data'' (p. 57). He asserts that ignoring these data effects may yield misleading interpretations of results, with major consequences for policy if the described central tendencies prove not to apply to certain subsets of locations. Finally, Hoogstra, Florax, and van Dijk (2005) reveal that the study results from U.S.- and non-U.S.-oriented studies are significantly different from each other, thereby confirming the idea of spatial nonstationarity in the jobs–people direction of causality, albeit at a broad geographical scale. Here, we reveal whether spatial heterogeneity also can be observed when geographical coverage is narrowed to a single province in the northern Netherlands by estimating the model with the complete data set as well as with four exclusive subsets of locations (spatial regimes). Specifically, a distinction is made between 347 settlements with fewer than 2,500 inhabitants, classified into the category of a ''small village'' (RPD 1999) and the remaining 45 ''urban'' settlements. Subsequently, the former category of ''rural'' settlements is divided into three subgroups, for which the respective population levels are 750–2,500 (102), 250–750 (118), and below 250 (127). The distinction is largely intuitively made to reflect coherent groups of locations and to ensure that each group contains a sufficient number of observations for comparison purposes.

The third data sampling issue investigated is the possible impact of the employment data. Employment is widely known to be extremely heterogeneous, comprising various subgroups that display different preferences for industrial location (Bollinger and Ihlanfeldt 2001). On the basis of the studies that have acted upon this premise by using sectoral employment data (e.g., Bollinger and Ihlanfeldt 1997; Schmitt et al. 2006), one is inclined to conclude that important group effects need to be considered in the jobs–people direction of causality. The lack of other specific studies, however, means that the effect of using data for alternative employment groups with varying labor intensity and consumer dependency has yet to be evaluated with rigorous statistical techniques. Therefore, here the selected population–employment interaction model is estimated by using data for total employment as well as for four private employment sectors: manufacturing; construction; retail; and the combination of finance, insurance, real estate, and services (FIRES). Together, these employment groups made up 51% of all full-time jobs for 2002 and 46% of the employment growth in the preceding 14-year period.[3]

Methodological issues

The literature offers a variety of suggestions for the methodological study of techniques that can be expected to influence the study results revealing population–employment interaction. Arguably the most obvious is measurement of the population and employment variables, with two techniques being in use that practically share the same amount of support, one using raw, unstandardized population and employment data, and one using standardized data by controlling for the area size of spatial units. However, as Mulligan, Vias, and Glavac (1999, p. 857) state: ''there is no a priori reason to expect that estimates based on levels will resemble estimates based on densities, as each approach represents an entirely different conceptualisation of the space-economy . . . .'' The findings from Glavac, Vias, and Mulligan (1998), in which levels as well as densities have been used, seem to indicate that alternative measurements indeed yield different conclusions as to the direction of causality. The findings by Hoogstra, Florax, and van Dijk (2005) appear to support this conjecture, suggesting furthermore that the nature of the effect also depends on a study's context.[4] On the basis of these arguments, a strong case exists for testing the model for both levels and densities and evaluating how exactly these different measurements change the results. Here, we go even further by distinguishing two alternative density measures, one in which population and employment are standardized by built-up area and one in which they are standardized by total area, in addition to using the unstandardized data. Referring to the ''different conceptualisations of the space-economy'' previously cited, the levels reflect the spatial distribution of employment and population (changes) as if it were a point pattern, whereas the densities by total area and built-up area depict the space-economy as a partitioned landscape of contiguous and noncontiguous area units, respectively. To compare growth across locations, the standardization by built-up area (which can be interpreted as the net size of a location) may be more appropriate than the routine standardization by total area (which corresponds to the gross size of a location). Especially in cases where locations show considerable variation in the part of the land that has not been built upon, and thus has not been used for industrial or household residential purposes (but mostly for agricultural purposes), the measurement by total land area may lead to completely different conclusions.

The second methodological issue investigated concerns the design of the spatial weights matrix *W*, which represents one of the most difficult and controversial aspects of a spatial econometric model.[5] As estimation results directly hinge upon the definition of the weight elements, a justification for the chosen specification of *W* is crucial. Stakhovych and Bijmolt (2009) provide Monte Carlo simulations for an array of different specifications of the weights matrix and corroborate evidence that weights matrices implying a high connectivity between spatial units are detrimental in finding the true underlying model and the mean square error of the estimated parameters. Existing population–employment interaction studies clearly lack a univocal specification of *W* (see Hoogstra 2011), meaning that there is no consensus regarding the type of weighting scheme that most realistically imposes structure on labor market relationships between regions. The findings by Boarnet, Chalermpong, and Geho (2005), in which the model described by equations (1a) and (1b) has been estimated using six alternative weight matrices, suggest that the model parameters are quite sensitive to different definitions of *W*.

Here we assess the crucial role attributed to weight matrix design by comparing the estimation results for three different weighting schemes. Specifically, the focus is on two standard weight matrices that dominate the literature—the fixed distance matrix and the inverse distance matrix—and one rare flow matrix based on commuting data, which has so far been used only in Boarnet, Chalermpong, and Geho (2005). The fixed and the inverse distance matrices are rather similar, and both are sparse matrices as typically the number of zero elements is rather high. The commuting flow matrix is a full matrix with, except for the diagonal elements, entries that are as a rule nonzero. Each of these matrices reflects different assumptions about the way in which spatial units of observations relate to each other and about how they are tied into larger labor market zones. The fixed distance matrix is a binary weighting scheme in which the matrix elements w_{ij} equal one for $d_{ij} \leq \delta$, and zero otherwise (where d_{ij} is the distance between locations i and j, and δ is a chosen distance threshold value). In this case, interaction is assumed to take place only between spatial units that are within a critical distance of each other (see, e.g., Henry, Schmitt, and Piguet 2001 for applications). Instead of matrix elements having values of zero or one, the weighted inverse distance matrix contains elements w_{ij} equal to $1/d_{ij}\alpha$ (with α denoting an a priori determined distance decay parameter).

By using such a matrix, labor market variables take the form of potential variables, with employment and population (growth) in nearby locations weighted more heavily (see, e.g., Boarnet 1992, 1994; Bollinger and Ihlanfeldt 1997). The flow matrix, with w_{ij} being a function of the number of commuters traveling between locations i and j, directly reflects the properties of labor market relations, and is, at least according to Boarnet, Chalermpong, and Geho (2005, p. 32), ''closer to a theoretical ideal of a commuter-shed than any other W matrix.''[6] In contrast to the other weighting schemes, the flow matrix does not impose a rigid spatial form on labor market zones of different locations, which in a regular spatial data arrangement all would be uniformly sized. Being based on real commuting data, the flow matrix is entirely flexible and allows each location to have a uniquely shaped labor market zone in ways that incorporate variations in commuting patterns across a region. The availability of an extremely rich data set of commuting patterns (see Van der Horn, Hellingwerf, and Veldhuis 2001) allows us to construct such a rare flow matrix. Based on these same data, we set the values for the threshold distance δ in the fixed distance matrix and the distance decay parameter α in the weighted inverse distance matrix to 6.8 (miles) and 0.92, respectively. The former approximates the average commuting distance, which is a standard criterion to determine the threshold value, whereas the latter is estimated from a spatial interaction model (see Hoogstra 2011 for details). By convention, all matrices are row standardized, which means that the row elements sum to one, although alternative coding schemes are available (see, e.g., Patuelli et al. 2006 for a more elaborate treatment).

The final, and arguably most important, issue to be investigated is that of model specification and estimation, as outlined earlier in ''Research design: econometric model and data.'' Because of the lack of an appropriate estimation technique until recently, the issue whether to include an SAR lag because of possible spatial dependence in the dependent variables has thus far been largely ignored (with the exception of Henry, Schmitt, and Piguet 2001; Carruthers and Mulligan 2008). The potential impact of failing to control for spatial dependence in the presence of such effects cannot be too strongly emphasized, as the parameter estimates revealing the jobs–people direction of causality then are biased and inconsistent. The wide divergence in inferences that typifies this literature is an outcome that one would typically expect when parameters are not properly estimated.[7] Therefore, to make a definite assessment about the impact of model specification and estimation, we perform three series of experiments: one with the original ''Boarnet model'' using 2SLS and two with the augmented Boarnet model, B-SAR, using both 2SLS and the GS2SLS estimator recently proposed by Kelejian and Prucha (2004). Importantly, the latter method adds two steps to a routine 2SLS procedure. First, the estimated disturbances u and v from the initial 2SLS estimations are used to estimate the autoregressive parameters, ρ and γ, in equations (2a) and (2b), respectively, by applying the generalized moments procedure described in Kelejian and Prucha (1999). Second, by applying a Cochrane–Orcutt-type transformation, the estimated autoregressive parameters are subsequently used to account for spatial dependence in the disturbances. Note that assessing the impact of model specification and estimation is somewhat different from that of the other study features being analyzed. Whereas those other features inform only about whether adopting a particular data sample or methodology affects estimation results, the issue of model specification and estimation addresses a more fundamental problem, namely, that of a possible inherent flaw, which makes a comparison of results infeasible.

Results

By changing the time period (6), region type (5), employment type (5), variable measurements (3), matrix design (3), and model specification and estimation (3), a total of $6 \times 5 \times 5 \times 3 \times 3 \times 3 = 4{,}050$ experiments were performed, generating a similar number of parameter estimates to be evaluated in the meta-analysis. Due to the different measurements of the population–employment relationship across these experiments, a comparison of the *magnitude* of the effects (as revealed by the size of the parameter estimates) is not permitted. Instead, these measurements only allow making inferences about the *sign* effects of α_4 and β_4. Accordingly, the analysis of study results necessarily takes the form of a vote-counting procedure in which the estimated sign and significance levels of α_4 and β_4 alone are used to determine whether the inferences from different experiments agree. Although such an evaluation is crude and puts considerable emphasis on statistical significance, it is intuitively very appealing because it seamlessly unites with the common practice in the literature of summarizing the estimation results by discrete categories. Here, the estimates for α_4 and β_4 are jointly used to discriminate between four categories of research findings, where 10% significance levels are used to determine whether or not these estimates differ from zero:

- (NI) no interaction (i.e., ''jobs do not follow people nor do people follow jobs''): α_4 and $\beta_4 \leq 0$,

- (JP) one-way causality running from population to employment (i.e., ''jobs follow people only''): $\alpha_4 \leq 0$ and $\beta_4 > 0$,
- (PJ) one-way causality running from employment to population (i.e., ''people follow jobs only''): $\alpha_4 > 0$ and $\beta_4 \leq 0$, and
- (DC) dual causality (i.e., ''jobs follow people and people follow jobs''): α_4 and $\beta_4 > 0$.

Considering that some of these estimates may be flawed because possible spatial dependence in the dependent variables is ignored, the ensuing discussion focuses on results not only from the entire set of estimations but also from the subset of GS2SLS-based estimations ($n = 1{,}350$) because they are known to be unbiased and consistent.

The last row in Table 1 reveals that most of the estimations (some three-quarters) fail to provide any evidence for population–employment interaction (either one-way or two-way), which is not particularly unusual for small area models of population–employment interaction (see Hoogstra, Florax, and van Dijk 2005). The remaining estimation results are spread over the three remaining categories that indicate a causal relation, with most of the results pointing toward PJ, closely followed by DC, and then JP. When we compare the distribution of results for the entire set of estimations with the subset of GS2SLS-based estimations, we may conclude that they are rather similar, although for the GS2SLS-based estimations slightly more results indicate the existence of a causal relation between population and employment, especially for both one-way causalities. Given the purposes of this study, the relationships with the underlying study characteristics are even more interesting than the differences in research findings per se. Table 1 furnishes an overview of the distribution of these findings across the four possible categories. Although the category ''no interaction'' contains the highest share for all study characteristics by far, substantial variation exists between the characteristics and between the percentages based on all estimations and those based only on GS2SLS.

Instead of discussing the results presented in Table 1, we prefer to discuss the differences between the study characteristics on the basis of a multivariate method that also gives insight into the statistical significance of the differences. Because the study results refer to four discrete categories, we adopt a multinomial logistic regression model, which reveals the influence of each of the study features on the likelihood of a categorical outcome, other things being equal (ceteris paribus). In our case, this model comprises three equations (a), (b), and (c) in which the respective dependent variables are defined as the log-odds that the estimation results indicate either JP, PJ, and DC, instead of no interaction (the reference alternative). From each group of study features that serve as explanatory variables, one category is omitted against which to compare. The estimated regression coefficients reveal the additive effect of each category compared with the omitted category (for which the coefficient is 0) and can be interpreted as the change in the log-odds. Intuitively more appealing is the interpretation of these coefficients as factors that indicate the

Table 1 Estimated Outcome by Characteristics, in Percentages, for All Estimates and the GS2SLS Estimates

Sample	All estimates ($n = 4{,}050$)				GS2SLS estimates ($n = 1{,}350$)			
Characteristic/outcome*	NI	JP	PJ	DC	NI	JP	PJ	DC
Time period								
1988–1992	74.8	4.1	11.3	9.8	70.2	9.3	10.7	9.8
1990–1994	80.6	7.1	3.7	8.6	81.8	6.7	3.6	8.0
1992–1996	85.3	5.2	5.3	4.1	82.2	6.2	8.0	3.6
1994–1998	73.6	6.4	11.6	8.4	68.9	9.8	13.8	7.6
1996–2000	65.3	6.5	15.7	12.4	64.0	8.0	15.1	12.9
1998–2002	65.6	8.1	12.9	13.3	60.9	10.2	14.2	14.7
Region type								
All regions	69.4	9.4	9.5	11.7	65.6	10.7	11.5	12.2
Urban ($\geq$ 2,500 inhabitants)	69.5	3.1	7.5	19.9	71.1	3.0	7.0	18.9
Rural (750–2,500)	77.3	2.6	17.5	2.6	76.3	4.1	16.7	3.0
Rural X (250–750)	82.5	6.3	7.3	4.0	76.3	10.7	9.3	3.7
Rural XX ($<$250)	72.5	9.9	8.5	9.1	67.4	13.3	10.0	9.3
Employment type								
Total employment	67.4	8.0	12.3	12.2	65.2	8.9	14.4	11.5
Manufacturing	83.6	3.8	8.4	4.2	79.3	5.2	10.0	5.6
Construction	73.3	6.4	10.4	9.9	71.1	7.4	12.2	9.3
Retail	71.1	4.3	9.1	15.4	69.3	7.0	8.5	15.2
FIRES	75.7	8.6	10.1	5.6	71.9	13.3	9.3	5.6
Variables measurement								
Levels	70.4	6.6	12.9	10.1	66.2	9.8	13.8	10.2
Density, built-up area	79.3	6.8	7.6	6.3	75.6	9.1	9.6	5.8
Density, total area	73.0	5.3	9.7	12.0	72.2	6.2	9.3	12.2
Weights matrix specification								
Fixed distance	77.6	5.7	7.8	8.9	74.9	7.6	9.8	7.8
Inverse distance	75.0	5.4	8.8	10.7	72.7	7.8	9.1	10.4
Flow	70.0	7.6	13.6	8.7	66.4	9.8	13.8	10.0
Model specification and estimation								
Boarnet (2SLS)	75.3	4.5	10.0	10.2				
B-SAR (2SLS)	76.1	5.9	9.3	8.7				
B-SAR (GS2SLS)	71.3	8.4	10.9	9.4				
Overall	74.2	6.2	10.1	9.5	71.3	8.4	10.9	9.4

*The labels for the different outcomes are as follows: NI, no interaction; JP, jobs follow people; PJ, people follow jobs; and DC, dual causality.

change in odds, which can be estimated by exponentiating these coefficients (i.e., taking the antilog with the base e). A positive coefficient means a factor is greater than one, thereby revealing an increase in the odds and hence implying a higher probability that this outcome occurs compared with the reference alternative. In contrast, a negative coefficient complies with a factor that is less than one, which means that the odds are decreased. In case a coefficient is not significantly different from zero, the factor equals one, which leaves the odds unchanged (for more details about the technique, see, e.g., Menard 2002), implying that for this particular study characteristic, the probability that this alternative occurs does not differ from the probability that the reference alternative occurs.

Table 2 reveals that the overall distribution of research findings mostly diverges across the content-related temporal, spatial, and sectoral employment categories, rather than across the experimental methodological issues. Specifically, the former group of study features reveals statistically significant estimates in each of the three metaregression equations, whereas these are noticeably absent for ''variables measurement'' in metaregression equation (a) and for most coefficients for the ''weights

Table 2 Metaregression Results, Multinomial Logit Using All Estimates ($n = 4{,}050$)

Logits†	(a) Logit (JP versus NI)			(b) Logit (PJ versus NI)			(c) Logit (DC versus NI)		
	b	exp(*b*)	prob.	*b*	exp(*b*)	prob.	*b*	exp(*b*)	prob.
Intercept	− 2.437		***	− 1.580		**	− 1.295		***
Time period (1988–1992)‡									
1990–1994	0.476	1.610	*	− 1.215	0.297	***	− 0.216	0.805	
1992–1996	0.090	1.094		− 0.903	0.405	***	− 1.053	0.349	***
1994–1998	0.457	1.580	*	0.044	1.045		− 0.144	0.866	
1996–2000	0.607	1.834	**	0.488	1.630	***	0.410	1.506	**
1998–2002	0.834	2.302	***	0.277	1.319		0.484	1.623	***
Region type (all regions)									
Urban	− 1.131	0.323	***	− 0.232	0.793		0.565	1.759	***
Rural	− 1.410	0.244	***	0.501	1.651	***	− 1.658	0.190	***
Rural X	− 0.596	0.551	***	− 0.483	0.617	***	− 1.323	0.266	***
Rural XX	0.004	1.004		− 0.166	0.847		− 0.313	0.731	*
Employment type (all employment)									
Manufacturing	− 0.990	0.372	***	− 0.643	0.526	***	− 1.373	0.253	***
Construction	− 0.320	0.726		− 0.277	0.758	*	− 0.325	0.722	*
Retail	− 0.690	0.502	***	− 0.375	0.688	**	0.201	1.222	
FIRES	− 0.048	0.953		− 0.338	0.713	**	− 0.973	0.378	***
Variables measurement (levels)									
Density, built-up	− 0.096	0.908		− 0.681	0.506	***	− 0.643	0.526	***
Density, total	− 0.256	0.774		− 0.338	0.713	***	0.156	1.169	
Weights matrix specification (fixed distance matrix)									
Inverse distance	− 0.017	0.983		0.169	1.184		0.250	1.284	*
Flow	0.411	1.509	**	0.701	2.016	***	0.100	1.105	
Model specification and estimation (Boarnet [2SLS])									
B-SAR (2SLS)	0.252	1.287		− 0.086	0.918		− 0.187	0.830	
B-SAR (GS2SLS)	0.692	1.998	***	0.148	1.159		− 0.033	0.968	

†See the note to Table 1 for the meaning of the labels. ‡Omitted categories are in parentheses. Critical significance levels are signaled by $^{*}<0.10$, $^{**}<0.05$, $^{***}<0.01$.

matrix'' and ''model specification and estimation.'' Also, for the content-related categories, the magnitude of the coefficients is larger, indicating a larger impact on the research findings. In metaregression equation (b), for example, the odds of finding ''people follow jobs'' instead of ''neither'' are lowered by 3.4 (= 1/0.297) when data from 1990–1994 rather than from 1988–1992 (reference category) are used. Likewise, in metaregression equation (a), examining ''urban'' and ''rural'' units rather than all spatial units decreases the odds of finding ''jobs follow people'' instead of ''neither'' by 4.1 (= 1/0.244) and 3.1 (= 1/0.323), respectively, whereas in metaregression equation (c) the change in odds (in this case of finding ''dual causality'' instead of ''neither'') due to examining these rural units is no less than 5.2 (= 1/0.190). Also in metaregression equation (c), the odds decrease by 4.0 when manufacturing employment data rather than all employment data are examined. By comparison, the change in odds related to model specification and estimation, variables measurement, and matrix design is never more than 2.0.

One by one, the different study features reveal some interesting findings. For ''time period,'' for example, the pattern observed is not clear-cut and thus hints at the influence of economic business cycles. Yet, the impression one obtains from metaregression equation (a) is conformity with the assumption of a shift toward a knowledge-based society in which the jobs follow people direction of causality is gaining significance over time (Florida 2002), followed by an increase in dual causality and a decrease over time in people follow jobs. With regard to the spatial aspect of urban versus rural, evidence that the odds of finding interaction, either one-way or two-way, is usually less when subsets of more homogeneous regions (rather than all data observations taken together) are being analyzed. This result also may reflect that most population–employment interactions take place between different categories of rural and urban regions rather than between similar types of regions. A notable exception is found for the urban and rural categories in metaregression equation (b) and (c), respectively, which arguably represent the most dynamic parts of the study region. Similarly, the odds appear to decrease when examining specific sectoral employment data and then manufacturing data especially. A possible explanation for the low interaction of population and employment in manufacturing might be that these industries are usually located in relatively large establishments on industrial sites that hardly change location. Van Dijk and Pellenbarg (2000) find empirical evidence supporting this contention and argue that the costs of moving for the industrial sector are generally higher because investment in capital stock and capital intensity is higher. People are also reluctant to reside near industrial activities. Therefore, the observed weak population–employment relationship appears to make sense, especially in view of the positive coefficient being observed for ''retail'' in metaregression equation (c), a sector for which this contention obviously does not apply.

With regard to the issue of variable measurement, the difference in results caused by standardizing the population and employment data is most telling when built-up areas (rather than total areas) are used as the basis for standardization, thereby negatively affecting the odds of finding people follow jobs and dual causality in particular. As for weight matrix design, the regular fixed distance and inverse distances weighting schemes give practically the same variation in research findings, whereas the unusual, but theoretically preferred, matrix based on commuting flows especially favors the finding of people follow jobs. Finally, with regard to the distinction between the two sets of 2SLS-based estimations, on the one hand, and the GS2SLS-based estimations, on the other hand, the latter is relatively strongly in favor of jobs follow people. Apparently, the inclusion of an SAR lag does not make a difference, as long as the model is not properly estimated by also taking into account the spatial dependence in the dependent variables (i.e., by using GS2SLS). Accordingly, we examine the relation between the GS2SLS-based estimations and the study characteristics in more detail.

Although the preceding assessment of the impact of the individual study features is done while controlling for the influence of model specification and estimation, it may crucially rest on a comparison of biased and inconsistent parameter estimates. The significant difference in parameter estimates observed between the 2SLS- and GS2SLS-based estimations suggests that this is true for several of the former estimations (because the variation in research findings would have been the same otherwise). Thus, to assess the true impact of the selected study features, the logistic regression analysis is repeated by solely using the subset of GS2SLS-based parameter estimates, which are known to be unbiased and consistent. Table 3 reveals that inferences based on this subset of estimations are somewhat different from those previously outlined. Specifically, many regression coefficients are no longer significantly different from zero at conventional statistical levels, especially those being associated with the time period and type of region. The signs of the regression coefficients are similar to those in Table 2, with the largest change in odds, within each study feature as well as across these features, being brought about by the same categories. Thus, while the estimation results appear to be particularly varied because, for example, data are used for different time periods, much of the variation can be ascribed to a bias in these results due to the use of an inappropriate estimator. This finding implies that the results obtained by using the GS2SLS estimator are less sensitive to variation in the study characteristics and thus give more reliable answers to the central question of this study with regard to the empirical nature of population–employment interaction.

Conclusions

The quasi-experimental meta-analysis summarized in this article includes a number of interesting findings. First, the various aspects of data sampling, variable measurement, and spatial weights matrix specifications are clearly secondary to the main issue of model specification and estimation. They are secondary because estimates and inferences are biased and inconsistent if spatial dependence exists in the dependent variables in addition to the right-hand-side endogenous variables.

Table 3 Metaregression Results, Multinomial Logit Using Only GS2SLS Estimates ($n = 1{,}350$)

Logits[†]	(a) Logit (JP versus NI)			(b) Logit (PJ versus NI)			(c) Logit (DC versus NI)		
	b	exp(*b*)	prob.	*b*	exp(*b*)	prob.	*b*	exp(*b*)	prob.
Intercept	− 1.427		***	− 1.180		***	− 1.398		***
Time Period (1988–92)[‡]									
1990–1994	− 0.509	0.601		− 1.274	0.280	***	− 0.369	0.692	
1992–1996	− 0.580	0.560		− 0.455	0.634		− 1.229	0.293	***
1994–1998	0.072	1.075		0.285	1.329		− 0.261	0.770	
1996–2000	− 0.066	0.937		0.452	1.572		0.400	1.491	
1998 – 2002	0.240	1.272		0.441	1.554		0.596	1.814	*
Region type (all regions)									
Urban	− 1.397	0.247	***	− 0.589	0.555	*	0.378	1.460	
Rural	− 1.142	0.320	***	0.203	1.225		− 1.628	0.196	***
Rural X	− 0.167	0.847		− 0.401	0.670		− 1.413	0.244	***
Rural XX	0.190	1.209		− 0.175	0.839		− 0.328	0.721	
Employment type (all employment)									
Manufacturing	− 0.765	0.465	**	− 0.600	0.549	**	− 0.992	0.371	***
Construction	− 0.283	0.754		− 0.271	0.763		− 0.330	0.719	
Retail	− 0.311	0.733		− 0.616	0.540	**	0.248	1.281	
FIRES	0.315	1.370		− 0.566	0.568	**	− 0.890	0.411	***
Variables measurement (levels)									
Density, built-up	− 0.218	0.804		− 0.525	0.591	**	− 0.762	0.467	***
Density, total	− 0.565	0.568	**	− 0.501	0.606	**	0.106	1.111	
Weights matrix specification (fixed distance matrix)									
Inverse distance	0.061	1.063		− 0.040	0.961		0.359	1.432	
Flow	0.399	1.491		0.489	1.631	**	0.407	1.503	

[†]See the note to Table 1 for the meaning of the labels. [‡]Omitted categories are in parentheses. Critical significance levels are signaled by $^{*}<0.10$, $^{**}<0.05$, $^{***}<0.01$.

Accordingly, the main methodological message from this article is that adding SAR lags offers an improvement to the regular Boarnet model. The methodology for estimating a B-SAR model is now available, thanks to Kelejian and Prucha (2004), and thus a reason for excluding these lags no longer exists.

Second, the subordinate nature of the other study features notwithstanding, these features reveal some significant impact on the findings of population–employment interaction. Specifically, the findings suggest that the parameter estimates are largely shaped by the region and time period under examination, and, equally important, employment group effects need to be considered when assessing the direction of causality. Also, the estimates appear rather sensitive to different measurements of a model's key variables, more so than to the application of alternative spatial weights matrices, which does not appear to be an issue with which future studies should be primarily concerned.

Overall, the results from this study suggest that findings about population–employment interaction alone are of little value if the impact of the underlying study features is not properly understood. For example, without understanding *why* the research findings are what they are, the potential for what is called ''value transfer'' (Florax, de Groot, and Heijungs 2002) remains remote. To illustrate this point, our knowledge still seems far from sufficient to predict the nature of causality for an unstudied site. This study shows that even when an analysis is restricted to a single province in the Netherlands, considerable spatial heterogeneity can be observed in estimates indicating population–employment interaction without having a clear understanding as to the reasons why. Having concluded that the estimates differ spatially, the next step is to understand why these differ by looking into the characteristics of the different locations in more detail.

Finally, the quasi-experimental meta-analysis proves to be a promising tool to assess the robustness of models to various implementation decisions. Staying with population–employment interaction models, one important area for further research seems to be the determination of whether, and if so how, a particular selection and combination of location-specific exogenous variables affects estimation results. Usually, variables are selected with a relatively weak justification from a set of ''obvious'' candidates for which data are readily available. Also, the focus of attention may be redirected toward estimation results for the model parameters that inform about the lagged adjustment process, an important issue that has largely gone unexplored (with the notable exceptions of Mulligan, Vias, and Glavac 1999; Boarnet, Chalermpong, and Geho 2005), and that has yet to be investigated using rigorous statistical techniques.

Acknowledgements

We gratefully acknowledge the comments and suggestions of Frank van Oort and two anonymous reviewers, and Guyslain Ngeleza's help with computer code. We thank the Province of Fryslân, the Netherlands Institute for Spatial Research (RPB), and Statistics Netherlands (CBS) for providing the data used in this study.

Notes

1 We refrain from discussing other aspects of the model, such as the underlying assumption of a lagged adjustment process because they fall outside the scope of this study. Readers interested in a more details are referred to, for instance, Carlino and Mills (1987), Boarnet (1992), and Mulligan, Vias, and Glavac (1999).

2 See Bollinger and Ihlanfeldt (1997, p. 185) for an intuitive reasoning behind this argument.

3 Individually, the employment sectors contribute as follows (given as percentages of total employment in 2002 and employment growth between 1988 and 2002, respectively): manufacturing, 19.7 and 12.0; construction, 9.5 and 11.2; retail, 8.0 and 6.5; and FIRES, 14.0 and 15.9. The sectors of education, government, and health care, in particular, make up the remainder of the full-time employment (growth), with agriculture being omitted from the analyses all together.

4 Reasons to use standardization techniques usually refer to the examination of spatial units that do not allow a straightforward comparison because of considerable differences in area size and/or population and employment size. Naturally, the effect of using standardization techniques seems to hinge on the combined characteristics of the spatial detail of an investigation and the region type under examination, which may differ between studies.

5 This is reflected by the variety of spatial weights matrix specifications that can be encountered across the spatial econometric literature (see, e.g., Anselin 2002 for an overview) and the series of studies about the misspecification of *W* (see, e.g., Florax and Nijkamp 2005, and the references herein).

6 The matrix is less than ideal compared with a simple distance-based matrix specification because the weights elements are less exogenous to the model. Maintaining the weights matrix as independent is important because the model otherwise becomes highly nonlinear with endogeneity that must be instrumented out. Typically, this is not the result one has in mind when designing a weights matrix (Anselin 2002). Notice that this study is not designed to draw inferences about which matrix specification is the most appropriate; rather, it is designed to determine whether the application of the different weights matrices yields different study results.

7 We thank an anonymous referee for pointing this out.

References

Anselin, L. (2002). ''Under the Hood: Issues in the Specification and Interpretation of Spatial Regression Models.'' *Agricultural Economics* 27, 247–67.

Banzhaf, H. S., and V. K. Smith. (2007). ''Meta-Analysis in Model Implementation: Choice Sets and the Valuation of Air Quality Improvements.'' *Journal of Applied Econometrics* 22, 1013–31.

Boarnet, M. G. (1992). ''Intra-Metropolitan Growth Patterns: The Nature and Causes of Population and Employment Changes within an Urban Area.'' Ph.D. Dissertation, Princeton University.

Boarnet, M. G. (1994). ''An Empirical Model of Intrametropolitan Population and Employment Growth.'' *Papers in Regional Science* 73, 135–52.

Boarnet, M. G., S. Chalermpong, and E. Geho. (2005). ''Specification Issues in Models of Population and Employment Growth.'' *Papers in Regional Science* 84, 21–46.

Bollinger, C. R., and K. R. Ihlanfeldt. (1997). ''The Impact of Rapid Rail Transit on Economic Development: The Case of Atlanta's MARTA.'' *Journal of Urban Economics* 42, 179–204.

Bollinger, C. R., and K. R. Ihlanfeldt. (2001). ''Spatial Interactions among Employment and Population Groups.'' Working Paper, University of Kentucky/Georgia State University.

Carlino, G. A., and E. S. Mills. (1987). ''The Determinants of County Growth.'' *Journal of Regional Science* 27, 39–54.

Carruthers, J. I., and G. F. Mulligan. (2007). ''Land Absorption in U.S. Metropolitan Areas: Estimates and Projections from Regional Adjustment Models.'' *Geographical Analysis* 39, 78–104.

Carruthers, J. I., and G. F. Mulligan. (2008). ''A Locational Analysis of Growth and Change in American Metropolitan Areas.'' *Papers in Regional Science* 87, 155–71.

Carruthers, J. I., and A. C. Vias. (2005). ''Urban, Suburban, and Exurban Sprawl in the Rocky Mountain West: Evidence from Regional Adjustment Models.'' *Journal of Regional Science* 45, 21–48.

Davidson, R., and J. G. MacKinnon. (1993). *Estimation and Inference in Econometrics*. Oxford: Oxford University Press.

Dubin, R. (2003). ''Robustness of Spatial Autocorrelation Specifications: Some Monte Carlo Evidence.'' *Journal of Regional Science* 4, 221–48.

Florax, R. J. G. M., and T. de Graaff. (2004). ''The Performance of Diagnostics Tests for Spatial Dependence in Linear Regression Models: A Meta-Analysis of Simulation Studies.'' In *Advances in Spatial Econometrics: Methodology, Tools and Applications*, 29–65, edited by L. Anselin, R. J. G. M. Florax, and S. J. Rey. Berlin: Springer Verlag.

Florax, R. J. G. M., H. L. F. de Groot, and R. Heijungs. (2002). ''The Empirical Economic Growth Literature.'' Discussion Paper TI 2002-040/3, Tinbergen Institute.

Florax, R. J. G. M., and P. Nijkamp. (2005). ''Misspecification in Linear Spatial Regression Models.'' In *Encyclopedia of Social Measurement*, 695–707, edited by K. Kempf-Leonard. San Diego: Academic Press.

Florida, R. (2002). *The Rise of the Creative Class: And How It's Transforming Work, Leisure, Community and Everyday Life*. New York: Basic Books.

Glavac, S. M., A. C. Vias, and G. F. Mulligan. (1998). ''Population and Employment Interactions in the Growth of United States Micropolitan Centers.'' *Urban Geography* 19, 632–56.

Henry, M. S., B. Schmitt, and V. Piguet. (2001). ''Spatial Econometric Models for Simultaneous Systems: Application to Rural Community Growth in France.'' *International Regional Science Review* 24, 171–93.

Hoogstra, G. J. (2011). ''The Location Changes of Jobs and People. Analyses on Causality, and Impacts of Geography and Gender.'' Doctoral Dissertation, Department of Geography, University of Groningen.

Hoogstra, G. J., R. J. G. M. Florax, and J. van Dijk. (2005). ''Do Jobs Follow People or People Follow Jobs? A Meta-Analysis of Carlino–Mills Studies.'' Working Paper, University of Groningen.

Kelejian, H. H., and I. R. Prucha. (1999). ''A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model.'' *International Economic Review* 40, 509–33.

Kelejian, H. H., and I. R. Prucha. (2004). ''Estimation of Simultaneous Systems of Spatially Interrelated Cross Sectional Equations.'' *Journal of Econometrics* 118, 27–50.

Knol, F. A. (1998). *Van hoog naar laag, van laag naar hoog: De sociaal-ruimtelijke ontwikkeling van wijken tussen 1971–1995*. Rijswijk: Sociaal en Cultureel Planbureau.

Menard, S. (2002). *Applied Logistic Regression Analysis*. Thousand Oaks: Sage Publications.

Mulligan, G. F., A. C. Vias, and S. M. Glavac. (1999). ''Initial Diagnostics of a Regional Adjustment Model.'' *Environment and Planning A* 31, 855–76.

Patuelli, T., D. Griffith, M. Tiefelsdorf, and P. Nijkamp. (2006). ''The Use of Spatial Filtering Techniques: The Spatial and Space-Time Structure of German Unemployment Data.'' Discussion Paper # 06-049/3, Tinbergen Institute, Amsterdam. Available at http://econpapers.repec.org/paper/dgruvatin/20060049.htm (accessed October 4, 2010).

Rey, S. J., and M. G. Boarnet. (2004). ''A Taxonomy of Spatial Econometric Models for Simultaneous Equations Systems.'' In *Advances in Spatial Econometrics: Methodology,* *Tools and Applications*, 90–120, edited by L. Anselin, R. J. G. M. Florax, and S. J. Rey. Berlin: Springer Verlag.

RPD. (1999). *Balans ruimtelijke kwaliteit*. Den Haag: Ministerie van Volkshuisvesting, Ruimtelijke Ordening en Milieubeheer.

RUG/CBS. (1999). *Regionale samenhang in Nederland: Bi-regionale input-output tabellen en aanboden gebruiktabellen voor de 12 provincies en de twee mainport regio's*. Groningen: Stichting Ruimtelijke Economie Groningen.

Schmitt, B., M. S. Henry, V. Piguet, and M. Hilal. (2006). ''Urban Growth Effects on Rural Population, Export and Service Employment: Evidence from Eastern France.'' *Annals of Regional Science* 40, 779–801.

Stakhovych, S., and T. Bijmolt. (2009). ''Specification of Spatial Models: A Simulation Study on Weights Matrices.'' *Papers in Regional Science* 88, 389–408.

Stanley, T. D. (2001). ''Wheat from Chaff: Meta-Analysis as Quantitative Literature Review.'' *Journal of Economic Perspectives* 15, 131–50.

Stanley, T. D., and S. Jarrell. (1989). ''Meta-Regression Analysis: A Quantitative Method of Literature Surveys.'' *Journal of Economic Surveys* 3, 161–70.

Van der Horn, A., R. S. Hellingwerf, and A. Veldhuis. (2001). *Atlas van de Friese Pendel. Pendel en woon-werk relaties in het jaar 2000*. Leeuwarden: Provincie Fryslân.

Van Dijk, J., and P. H. Pellenbarg. (2000). ''Firm Relocation Decisions in The Netherlands: An Ordered Logit Approach.'' *Papers in Regional Science* 79, 191–219.

Vias, A. C., and J. I. Carruthers. (2005). ''Regional Development and Land Use Change in the Rocky Mountain West.'' *Growth and Change* 36, 244–72.