Do ‘jobs follow people’ or ‘people follow jobs’?
A meta-analysis of Carlino-Mills studies

Gerke J. Hoogstra*, Raymond J.G.M. Florax**, and Jouke van Dijk*

Paper prepared for the 45th Congress of the European Regional Science Association
23-27 August 2005, Amsterdam, the Netherlands

draft (do not quote without permission of the authors)

Abstract

The issue whether regional development can be associated with population driving employment changes or employment driving population changes (do ‘jobs follow people’ or ‘people follow jobs’?) has recently attracted considerable interest. Much of the interest herein stems from alleged inconsistencies in the empirical evidence, which naturally raises questions as for the reasons why. Arguably, the nature of causality differs across space as well as time, while speculations have been rife about a number of methodological issues that may play a crucial role in shaping the research outcomes. In this paper a preliminary attempt is described in clarifying these matters, by focusing on an articulate literature of 37 so-called ‘Carlino-Mills studies’. Specifically, a statistically supported literature review, referred to as ‘meta-analysis’, is provided in which the study results are evaluated and systematically related to a variety of underlying study characteristics. By listing 308 study results reported in this literature, it is revealed that the empirical evidence is conform popular belief highly inconclusive, albeit that most results point towards ‘jobs follow people’. The findings of the meta-regression analysis indicate that the spatial characteristics of the data, model specification, and variables measurement in particular affect the research outcomes that indicate the jobs-people direction of causality. No evidence is found that the examination of data referring to a particular time period, population and/ or employment group make much of a difference.

Keywords: population, employment, causality, meta-analysis

*Urban and Regional Studies Institute (URSI) Faculty of Spatial Sciences University of Groningen, the Netherlands; **Department of Agricultural Economics, Purdue University, United States and Department of Spatial Economics, Free University of Amsterdam, the Netherlands. Corresponding author: Jouke van Dijk (j.van.dijk@rug.nl)

1. Introduction

With significant changes in the spatial landscape taking place, there has recently been substantial interest in the essential location dynamics of residences and employment, and the nature of their relationship in particular. The latter is echoed in the classic phrase questioning whether 'jobs follow people' or 'people follow jobs', which reflects a longstanding debate in urban and regional science. Alleged striking discrepancies in the empirical evidence for the jobs-people direction of causality, as pointed out by qualitative literature reviews (see, e.g., Bollinger and Ihlanfeldt 2001; Sohn and Hewings 2000) mean that this debate is far from being settled. Moreover, it has evoked various speculations about the possible reasons for the differences in empirical findings, as illustrated by Carruthers and Vias (2003) who recently suggested that "*the character of the process probably varies from region to region and maybe even from time period to time period*" (p. 4). Alternatively, as pointed out by Boarnet *et al*. (2002), among others, it may also be assumed that the mixed empirical evidence represents a scientific artefact, which stems from methodological differences between the studies.

Thus far, no efforts have been undertaken to make a precise statement about the actual variation in research findings, which suggests that the reputed mixed empirical evidence for the jobs-people direction of causality is viewed upon as a 'stylised fact' that needs no further validation. Alternatively, researchers may have refrained from comparing the 'unique' research findings from the individual studies because of considerable heterogeneity pertaining to the data and methodologies used in these studies, which make them seemingly not amenable to summary. Whatever reason, the absence of an integrated research review is far from ideal. Most importantly, it remains undecided which factors can be held responsible for the alleged variation in research findings and, consequently, leaves unanswered *why* the research findings are *what they are*.

Considering the above, the objective of this paper is to comprehensively synthesise the literature in which the 'jobs follow people' and 'people follow jobs' hypotheses have been empirically tested and to evaluate its main findings. Specifically, we concentrate on an articulate body of studies, which in succession to Carlino and Mills (1987) have adopted a simultaneous equations model with adjustment lags. The widespread use of this model, mounting up to 37 studies in the 1987-2003 period, calls for a research synthesis to take stock of what is currently known about population-employment interaction. For the evaluation of this literature, we adopt a quantitative approach referred to as 'meta-analysis', which constitutes the application of statistical techniques to collections of empirical findings from previous studies for the purpose of integrating, synthesising, and making sense of them (Wolf 1986). Among these techniques is a meta-regression analysis, which is well suited to clarify seeming inconsistencies in study results across a literature. By listing the study results encountered across the Carlino-Mills studies and relating them to the characteristics of the data being used, we attempt to shed light on the validity of the assumption that substantive differences between regions and time periods are foremost responsible for the ambiguity surrounding the issue of causality, which is of pivotal importance for theory as well as for policy considerations.

Furthermore, we assess the robustness of the study results against particular methodological study features in order to identify the factors that induce the greatest bias or distortion in the results, and which should be carefully selected in future primary research accordingly.

The outline of this paper is as follows. Section 2 discusses meta-analysis as a complementary approach to summarising and synthesising research findings. Section 3 introduces the econometric framework that underlies all models developed in the spirit of Carlino and Mills. Subsequently, it confers the selection of studies that pertain to this literature and describes the variation among these studies as for the study results and study characteristics. Section 4 presents the set-up and results of the meta-regression analysis in which the variation in research findings across the Carlino-Mills literature is systematically explained. Finally, Section 5 recapitulates the main findings of this study and suggests avenues for further research.

2. Meta-analysis for research synthesis

The research literature, as it is often pointed out, is growing at an exponential rate. The latest 'subject and author index' published by the *International Regional Science Review*, for instance, reveals that nowadays a staggering number of sixty-six regional science and related journals can be discerned. For comparison, a mere eighteen studies were compiled for the first issue of this index 25 years before. Likewise, Suriñach *et al*. (2003) have shown that in the nine leading international regional and urban oriented journals alone the number of published papers has risen by more than 30 percent between 1990 and 2000. As research results accumulate, it becomes increasingly complicated to make sense of the flood of information. Researchers normally do not aspire to replicate or re-analyse, but typically pursue the new, the novel or, at the least, they attempt to extend what is considered to be the current state of knowledge. With no two studies being exactly alike, it is difficult to determine whether the differences between the study outcomes are attributable to methodological, contextual, or substantive variations in the research studies by using informal methods of narrative review that allow these issues to remain easily undetected (Rudner *et al*. 2002).

Meta-analysis is a quantitative literature review that is tailor-made to compare research findings obtained in different studies. Being introduced by Glass (1976) in the mid-1970s, meta-analysis is best seen as a statistical approach towards reviewing and summarising the literature (Stanley 2001). It complements the casual, narrative discussions of research studies that typify traditional attempts to make sense of the rapidly expanding research literature. Although ordinary literature reviews are valuable in their own right, there are a number of disadvantages in solely relying on such surveys (Dalhuisen *et al*. 2003). For instance, they can usually be criticised for a lack of objectivity in the selection of studies, which makes the comparison of study results largely arbitrary (Van den Bergh *et al*. 1997). Although one could argue that alternative methods of research synthesis are not necessarily free from subjectivity either, the selection procedure followed in a meta-analysis has to be explicit and is therefore more transparent (Florax *et al*. 2002). Next, qualitative literature surveys generally rely on some sort of vote-counting procedure, which is not very powerful in coming up with the right conclusion. Hedges and Olkin (1980) have shown that this technique, which essentially boils down to simply tallying significant results of a specific sign and non-zero results, contains the basic flaw that it tends to make the wrong interference when the number of studies increases.[1] In addition, the crudity of vote counting by looking solely at the sign-effects leaves much to be desired. Statistical significance alone is insufficient to determine whether the results of different studies agree (Hedges 1997). The most elemental problem associated with qualitative review techniques, however, is that they are not equipped to cope with the complexity of a literature, in which many factors are operating in a non-isolated way and interconnected to each other through relationships that can only be identified in a mathematical framework. In words of Rudner *et al.* (2002, p. 2): "*Confronted with the results of 20 similar studies, the mind copes only with great difficulty. Confronted with 200, the mind reels. Yet that is exactly the scope of the problem faced by a researcher attempting to integrate the results from a large number of studies*". Meta-analysis helps the reviewer to overcome such problem. In a meta-analysis, hundreds of research studies can be coded and interpreted using statistical methods similar to those applied in an individual empirically designed study. The quantitative orientation implies that the studies are compared in a systematic way that is more objective and exact than a narrative review. Moreover, given its statistical nature, meta-analysis furnishes more insight and greater explanatory power than the mere listing of studies and research findings (Rudner *et al.* 2002). As for the use of statistical techniques, meta-analysis is not unlike primary and secondary analysis. It takes a different stance with regard to the data that are investigated. Whereas primary and secondary analysis can be referred to as an original and an extended examination of a single dataset, respectively, meta-analysis uses aggregate data from existing studies and thus exploits a number of datasets (Glass 1976). It may as a result arrive at conclusions that are not available to primary or secondary analysis. For instance, individual studies provide relatively good estimates of the sampling uncertainty of results, but usually rather poor estimates of the consequences of non-sampling issues, such as research design, model specification, and estimation technique (Hedges 1997). Meta-analysis opens the possibility of investigating these non-sampling issues, which are usually constant within a study, in a multivariate framework that allows the assessment of marginal effects (Florax *et al.* 2002).

The basic procedures followed in studies that resort to meta-analysis for research synthesis comprise four different stages, which are not dissimilar to the various steps undertaken in primary research (Cooper and Hedges 1994; Stanley 2001). The first stage relates to the *problem formulation*, which includes the definition of the research question to be summarised and the identification of a research design that guides study sampling and data collection. The second stage concerns the *data collection*, in which the literature is searched upon relevant studies that meet specified criteria for inclusion. The third stage involves the *data evaluation*, which includes the extraction of those bits of information that help to answer the question that impels the research. At

[1] This is because the Type-II statistical errors (i.e., failing to detect a true effect) in the original studies do not cancel out (see also Hedges and Olkin 1985).

this stage, a database is constructed for which study characteristics are indexed and coded according to the objectives of the review and as checks on threats to validity. Additionally, the study outcomes are transformed to a common metric so that they can be compared. In the final stage of *analysis and interpretation* statistical techniques are applied to discover the consistencies and to account for the variability in the studies that make up the database. Among the array of statistical techniques that pertain to the toolbox of meta-analysis (see e.g., Cooper and Hedges 1994 for an overview), meta-regression analysis is generally seen as one of the most powerful (Stanley and Jarrell 1989). In contrast to standard meta-analysis, which combines effect sizes of different studies in one overall effect size, meta-regression is especially designed to assess the variation in the effect size against one or more of the underlying study characteristics.[2] The general framework underlying a meta-regression analysis reads as:

$$(1)\quad Y_i = \alpha + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \ldots + \beta_K X_{i,K} + \varepsilon_i$$

Above, Y_i is the effect size of study i, $X_{i,k}$ are (k = 1,2,…, K) explanatory variables that portray different features associated with study i, α and β_k are unknown model parameters to be estimated, and ε_i is the random disturbance term of study i. The model described by Equation (1) is extremely suited for the evaluation of a literature that is regarded as being indecisive, with no clear or consistent answers given to the questions being addressed. In this respect, the Carlino-Mills literature that questions the jobs-people direction of causality, and for which it is widely assumed that a hodgepodge of conflicting research findings can be observed, provides an excellent case. The model allows testing specific reasons why these findings may appear contradictory or overly varied, including several issues that relate to model specification and data selection, of which some have already been the subject of considerable speculation.[3]

3. The Carlino-Mills literature

3.1 Study sampling

Following the various stages of conducting a meta-analysis as described above, we start with a discussion of the specificities of the Carlino-Mills model, which helps to retrieve relevant studies and to distinguish them upon their main characteristics. The central feature of this model is the combination of a simultaneous equations system with adjustment lags, which has become the standard to investigate population-employment interaction. The basic properties of the model are given by Equations (2a) and (2b), which reveal that the population and employment equations

[2] An ‘effect size’ refers to a quantitative measure of the study output, such as a standardised mean difference, regression coefficient, odd ratio, elasticity estimate, or a similar statistical summary indicator, which serves as input for the meta-analysis.

[3] In fact, the robustness of estimation results (or the lack of) along different dimensions of research has been made an important issue of investigation in several Carlino-Mills studies (Mulligan *et al.* 1999, Boarnet *et al*. 2002, and Boarnet and Chalermpong 2002), albeit without the application of formal meta-analysis techniques.

enclose on their right-hand side (RHS) the dependent variables of both equations. Herewith, the dependent variable from the own equation appears as an explanatory variable in a time lagged form (and is therefore predetermined), whereas the dependent variable from the opposite equation appears as an endogenous explanatory variable. The former exemplifies the lagged adjustment process that is assumed, and allows making inferences about the speed of adjustment by households and firms, whereas the latter is the essential feature of a simultaneous equation system, in which the endogenous variables are jointly determined. Below, the Carlino-Mills model is only described in its most elementary form. Besides the lagged dependent variables that appear as explanatory variables it may include other exogenous and endogenous variables (for details about the theoretical underpinnings, as well as the estimation procedures see, e.g., Boarnet 1992; Mulligan *et al*. 1999; Hoogstra 2005).

(2a) $\tilde{P}_t = \mathrm{a}_0 + \mathrm{a}_1 P_{t-1} + \mathrm{a}_2 (I + \tilde{W})\tilde{E}_t + \ldots + \mathrm{u}_t$

(2b) $\tilde{E}_t = \mathrm{b}_0 + \mathrm{b}_1 E_{t-1} + \mathrm{b}_2 (I + \tilde{W})\tilde{P}_t + \ldots + \mathrm{v}_t$

(2c) $\tilde{P}_t = P_t - \delta_1 P_{t-1}$

(2d) $\tilde{E}_t = E_t - \delta_2 E_{t-1}$

(2e) $\tilde{W} = \delta_3 W$

Where P_t (P_{t-1}) is an n by 1 vector of population levels at time t (time t-1), E_t (E_{t-1}) is an n by 1 vector of employment levels at time t (time t-1), I is an n by n identity matrix, W is an n by n weights matrix, δ_1, δ_2, and δ_3 are scalars that are either 0 or 1, a and b are model parameters to be estimated, and u and v are n by 1 vectors with the random disturbance terms

Table 1 *Taxonomy of Carlino-Mills models*

	LHS δ_1, δ_2	**RHS** δ_1, δ_2	δ_3	**Model**
1	0	0	0	Carlino and Mills (1987)
2	0	0	1	Deitz (1998)
3	1	0	0	Mills and Carlino (1989)
4	1	1	0	Boarnet (1992)
5	1	1	1	Boarnet (1992)

For the meaning of scalars δ_1, δ_2, and δ_3 see above.

Over the years, various Carlino-Mills model specifications have been adopted, which can be broadly differentiated upon the values for the scalars δ_1 and δ_2 in Equations (2c) and (2d), which reflect the endogenous population and employment variables either as changes or end-of-period levels, and δ_3 in Equation (2e) which reveal whether the explanatory endogenous variables are recomputed in conjunction with a spatial weights matrix, in order to control for spillover effects between locations or not. For instance, the combination of $\delta_1 = \delta_2 = 0$ and $\delta_3 = 0$ corresponds to the standard framework used by Carlino and Mills (1987), which measures the endogenous variables as levels (instead of changes) and which the RHS variable is not spatially weighted (see Table 1). The 'Boarnet model' in which these variables are measured as changes, i.e., $\delta_1 = \delta_2 = 1$, and the RHS endogenous variable is spatially weighted, i.e., $\delta_3 = 1$, appears on the opposite side of the spectrum of model specifications. In between, there is the modified non-spatial Carlino-Mills framework that portrays the LHS variables as changes (in contrast to the RHS endogenous variables that appear as end-of-period levels). Two uncommon specifications complete the list of alternative Carlino-Mills model specifications in Table 1. One is a framework introduced by Deitz (1998) that controls for possible spill-over effects by adopting an endogenous variable on the RHS of each equation that is spatially weighted (similar to that of Boarnet), but with all the endogenous variables appearing in the form of end-of-period levels (akin to the regular Carlino and Mills model). The other is an altered version of the Boarnet model in which the cross-regressive spatial lag (the part described by the multiplication with W) is no longer part of the RHS endogenous variable, but in which it appears as a separate additional explanatory variable (see Boarnet 1992) or in which it is omitted all together (see Bao 1996; Schmitt *et al*. 1999; Henry *et al*. 2001).

Using Equations (2a) and (2b) as a guideline to make a precise decision as to which studies to include in the meta-analysis (so to ensure that the literature under examination is sufficiently homogenous), a retrieval procedure was started to compile all relevant documents. As the thoroughness and completeness of a literature retrieval is a crucial factor in determining the validity and the extent to which results of a meta-analysis can be generalised (Cooper and Hedges 1994), a number of different retrieval methods were employed, which included browsing bibliographic databases like *EconLit* (http://www.econlit.org) and *ProQuest* (http://www.umi.com), chasing references via the *Social Sciences Citation Index* (http://www.isinet.com/isi/products/citation/ssci), consulting authors, and using the *Google* search engine (http://www.google.com). Moreover, recent conference programs of the North American and European supraregions of the *Regional Science Association International* (RSAI) were screened for relevant paper presentations. Ultimately, the different search strategies identified 37 studies that met the specified criteria for inclusion and which agreed to the quantification of the study characteristics and study results for the purposes of comparison in a meta-analysis. The rather strict definition applied meant that a few similar appearing studies (McDonald 1989; Schmitt 2000; Pantuosco *et al*. 2001) in which a deviating adjustment process was assumed (as reflected in the lagged dependent variables being measured as lagged changes instead of base years levels) were left out of consideration. Besides, several documents in which the coefficients of the essential endogenous variables were either not estimated (Deller *et al*. 2001) or did not permit a straightforward interpretation (Henry *et al.* 1999) were discarded. Finally, several studies could not be obtained (Holmberg and Strömqvist 1988) or were not completed in time (Carruthers and Mulligan 2004; Nzaku and Bukenya 2004; York and Munroe 2004) to be included in the meta-analysis.

3.2 Study results

The studies that constitute the database for the meta-analysis can be evaluated upon the estimates for the regression coefficients a_2 and b_2 in Equations (2a) and (2b), which reveal the effect of employment on population and population on employment, respectively. Because of the different ways in which the population-employment relationships have been measured across the Carlino-Mills literature (discussed in detail in the next paragraph) the estimation results do not permit a comparison of the *magnitude* of the effects (as revealed by the size of the coefficients). Instead, the estimated coefficients only permit making inferences about the *sign* effects of a_2 and b_2 obtained in the different studies. Accordingly, the analysis of the study results necessarily takes the form of a vote-counting procedure in which the estimated sign and significance levels of a_2 and b_2 alone are used to determine whether or not the results of the studies agree. As already discussed in Section 2 such a method is rather crude and puts too much emphasis on statistical significance, given that the *economic* significance in terms of the size of the estimated effects is ultimately of overriding importance (McCloskey 1985). Notwithstanding these drawbacks, the vote-counting procedure is intuitively very appealing since it seamlessly unites with the common practice in the literature to narrow down the discussion of causality to the simple questions whether or not ‘jobs follow people’ and/ or ‘people follow jobs’. For the purposes of comparison in a meta-analysis the estimates for a_2 and b_2 are combined to obtain four categories of research findings:

- I: Both a_2 and b_2 are not significant at conventional statistical levels or do not display the theoretically expected positive sign, i.e., ‘jobs do not follow people nor do people follow jobs’.
- II: Only b_2 is positive and statistically significant, suggestive of uni-directional causality running from population to employment, i.e., ‘jobs follow people’.
- III: Only a_2 is positive and statistically significant, suggestive of uni-directional causality running from employment to population, i.e., ‘people follow jobs’.
- IV: Both a_2 and b_2 are positive and statistically significant, suggestive of dual or bi-directional causality, i.e., ‘jobs follow people and people follow jobs’.

From the selected 37 studies that pertain to the Carlino-Mills literature 308 study results can be derived that reveal the character of population-employment interaction in accordance with the categorisation made above. To avoid biases from certain results being ‘overrepresented’ through double counting, the compilation of study results comprises only those that are ‘exclusive’. This implies that at least one of the underlying study features that can be discerned is different from the other study results included in the sample. From Table 2, which reveals the distribution of the study results over the four abovementioned categories, it can be concluded that the empirical evidence for the jobs-people direction of causality is conform popular belief extremely mixed.[4] As well as

[4] Here, 10-percent statistical levels have been used to determine whether the coefficients found in the literature are significantly different from zero, which is the standard criterion adopted in most studies, with several studies (among which Mulligan *et al*. 1999) only informing about the significance of their estimates at this level.

between studies there is proof of substantial variation in research outcomes *within* studies, which prevents drawing clear-cut inferences with regard to the nature of population-employment interaction. In fact, about all the 26 studies that provided multiple study results (with the exception of Bao *et al*. 1999; Vergolino and Jatobá 2001; Rosenberger *et al*. 2002) have come up with contradictory findings, hence inevitably leading to questions as for the reasons why.

The final row of Table 2, which adds up the study results from the individual studies, shows that most of the evidence is pointing towards 'jobs follow people', but that the number of study results supporting the 'people follow jobs' hypothesis is only slightly less. It can also be seen that the number of findings that indicate the absence of any form of causality does not lag far behind and that the category with the lowest number of observations (pointing towards dual causality) still represents about one-fifth of all study results. By calculating a cross tabulation between the separate findings for a_2 and b_2 it can be seen from Table 3 that the distribution over the rows is significantly different from that over the columns ($\chi^2 = 11.530$, Cramer's V = 0.193, p = 0.000).[5] Specifically, it appears that the Carlino-Mills studies tend to yield contrasting outcomes for a_2 and b_2. It applies to both coefficients that the number of study results revealing the absence of a positive causal relationship exceeds the number of estimates revealing the presence of such a relationship, albeit with ratios of 53-47 and 57-43, respectively, slightly less so for b_2 than for a_2. The overall picture, though, seems largely shaped by the studies that provide numerous estimation results. Notably, the study of Mulligan *et al*. (1999), which appears to be strongly leaning towards 'people follow jobs', contributes with no less than 150 study results to nearly half of the observations. In addition to Mulligan *et al*. (1999), there are several other studies that add numerous estimation results, which make it difficult to weight up the evidence with regard to the nature of causality. Herewith, several of the studies included in the sample are basically part of one and the same research project, being based on the same dataset and with the same authors involved. To account for the multiple measurements within single studies and across associated studies the distribution over the four categories is in Table 3 also given for a weighted sample in which these measurements are treated as *independent weighted replications* (Bijmolt and Pieters 2001). Specifically, with 22 seemingly independent clusters of studies being distinguished (see Table 2 for details) the 308 observations are given weights that count up to 14 per cluster, meaning that the different research projects contribute equally to the analysis (for example, the 150 observations from Mulligan *et al*. 1999 are each assigned a weight of 0.093). It is revealed that the *associated* study results play a crucial role in evaluating the evidence for the jobs-people direction of causality. With the study results being weighted, the observed and expected distributions over the categories are about similar and the outcomes for a_2 and b_2 are no longer significantly different from each other ($\chi^2 = 0.160$, Cramer's V

[5] Pearson's Chi-square test compares the observed distribution with the expected distribution over the cells to conclude whether or not there is an association between the row and column variables. Herewith, a large chi-square statistic corresponds to a small p-value and the null hypothesis of no association is rejected if the p-value is small enough (say < 0.05). Cramer's V is a chi-square-based measure of nominal association that gauges the strength of the relationship (for which the upper bound is 1).

= 0.023, p = 0.689). The weighting of the study results brings about that the share of findings indicating ‘people follow jobs’ decreases considerably to the advantage of ‘dual causality’ and ‘jobs follow people’ in particular. With a share of 45.5% the latter category is now about twice as large as the category of ‘no interaction’, which, despite becoming the second largest category (practically together with ‘dual causality’), has seen its share being reduced to 21.8%. With a ratio of 33-67 the number of study results pointing towards employment driving population changes ($a_2 > 0$) as opposed to employment *not* driving population changes ($a_2 \leq 0$) has become highly disproportionate in favour of the latter. The corresponding ratio with regard to the coefficient in the employment equation has become equally disproportionate, but in contrast to a_2 in the population equation, strongly in favour of finding a positive and statistically significant estimate for b_2.

From Tables 2 and 3 it appears that the empirical evidence for the nature of population-employment interaction strongly depends on the particular set of studies under examination. In this respect, the conclusion that the studies most frequently point towards ‘jobs follow people’ is alone of little value and requires additional insight in the characteristics of the underlying studies to be judged by its true merits. The finding of one-way interaction running from population to employment (i.e., ‘jobs follow people’) may dominate the literature because of most studies being, for instance, United-States oriented. Similarly, there may be a considerable bias in the sample of studies towards the use of data that relate to a specific time period. Hence, in order to understand the study results above, the information from each study about the particulars of its data sets, methods, research design, and the like must be identified and coded. The salient study characteristics that can be expected the make most of the differences are to be discussed in detail in the next paragraph, before Section 4 presents the estimation results of a meta-regression model by which the impact of these features is ascertained.

Table 2 *Overview study results across the Carlino-Mills literature*

	I	II	III	IV	*n*	
Carlino and Mills (1987)	0	0	1	1	2	●
Mills and Carlino (1989)	1	0	0	1	2	●
Danielson and Wolpert (1991)	1	0	0	0	1	
Boarnet (1992)	3	0	0	5	8	○
Boarnet (1994a)	0	1	0	0	1	○
Boarnet (1994b)	0	1	0	0	1	○
Luce (1994)	0	1	0	0	1	
Mills and Lubuele (1995)	0	0	0	1	1	
Bao (1996)	2	5	1	0	8	○○
Clark and Murphy (1996)	0	1	0	1	2	
Bollinger and Ihlanfeldt (1997)	8	10	1	1	20	
Duffy-Deno (1997a)	0	0	1	0	1	*
Duffy-Deno (1997b)	0	1	0	0	1	*
Henry *et al*. (1997)	1	0	0	0	1	○○
Kristensen and Henry (1997)	0	1	0	0	1	
Barkley *et al*. (1998)	1	0	1	0	2	○○
Deitz (1998)	2	6	0	0	8	
Duffy-Deno (1998)	0	2	2	0	4	*
Glavac *et al*. (1998)	0	1	1	0	2	

Vias (1998)	2	10	0	3	15	**
Bao *et al*. (1999)	2	0	0	0	2	○○
Mulligan *et al*. (1999)	28	37	66	19	150	
Schmitt *et al*. (1999)	0	2	1	2	5	●●
Vias and Mulligan (1999)	0	1	0	0	1	**
Schmitt and Henry (2000)	1	0	2	1	4	●●
Schmitt *et al*. (2000)	4	0	2	0	6	●●
Argo (2001)	2	1	0	0	3	
Henry *et al*. (2001)	0	3	1	1	5	●●
Holmberg *et al*. (2001)	0	0	0	1	1	
Vergolino and Jatobá (2001)	0	2	0	0	2	
Arauzo-Carod (2002)	1	2	2	5	10	
Boarnet and Chalermpong (2002)	8	1	1	0	10	○●
Boarnet *et al*. (2002)	9	1	1	1	12	○●
Rosenberger *et al*. (2002)	0	3	0	0	3	
Schmitt *et al*. (2002)	2	1	1	0	4	●●
Carruthers and Vias (2003)	1	0	0	4	5	**
Edmiston (2003)	0	3	0	0	3	
Total	79	97	85	47	308	

I = no interaction; II = jobs follow people; III = people follow jobs; IV = dual causality
●; ○; ○○; *; **; ●●; ○● denote related studies being part of the same research project

Table 3 *Distribution of study results across alternative samples (in %)*

			$B_2 \leq 0$			$B_2 > 0$		*Total*
$A_2 \leq 0$	(a)	**I**	25.6	(30.4)	**II**	31.5	(26.7)	57.1
	(b)		21.8	(22.3)		45.5	(44.9)	67.2
$A_2 > 0$	(a)	**III**	27.6	(22.9)	**IV**	15.3	(20.0)	42.9
	(b)		11.4	(10.8)		21.4	(21.9)	32.8
Total	(a)		53.2			46.8		
	(b)		33.1			66.9		

(a) unweighted sample of study results; (b) weighted sample of study results (n = 308)
See Table 2 for the meaning of I, II, III, and IV. In parentheses the expected distribution, calculated by dividing the products of the row and column totals by the grand total.

3.3 Study descriptors

As for a sample of studies for which it can realistically be assumed that the study results are determined by the nature of the substantive phenomenon under investigation and the nature of the methods used to study it (like the Carlino-Mills literature), three broad categories of study descriptors can be distinguished (Cooper and Hedges 1994). The first, and the most important class of study descriptors, at least from a theoretical perspective, is the set of features substantively pertinent to characterising the issue that prompts the investigation. Among such matters in the Carlino-Mills literature are the data characteristics that concern the geographical and temporal setting, and the population and employment types. It is these *substantive* study characteristics that need to be examined in order to derive conclusions as to whether a population-employment relationship holds for certain groups and not for others, and whether or not there is a developmental sequence that yields different results for different regions and time periods. The second class of study descriptors, not being related to substantive aspects of the phenomenon under examination, involves possible sources of distortion, bias, or artefact in the study results. The study descriptors of this sort represent the methodological or procedural aspects of the manner in which the primary studies have been conducted. These include variations in research design, model specification, variables measurement and the like, which might yield different study results even if all studies were investigating exactly the same dataset. The analyses of the relationships between the *methodological* study characteristics and the study results provide useful insights with regard to the research methods that produce similar results, thereby satisfying the criterion of 'convergent validity' (Summers and MacKay 1976), and the methods that make any difference and, hence, should be carefully selected by researchers. The third category of study descriptors consists of the factors *extrinsic* to both the substantive phenomenon and the research methods that are assumed to affect the study results. These include the characteristics of the researcher (e.g., gender, disciplinary affiliation), research context (e.g., number of researchers involved), and reporting (e.g., form of publication), which are not believed to directly shape study results, but nonetheless may be related with the results and thus need to be controlled for.

To highlight potential sources for the variation in research findings a database is constructed in which the Carlino-Mills studies are coded according to a variety of study descriptors, corresponding to the broad categorisation made above. While a plethora of study dimensions can be discerned, we concentrate the discussion in this paper on a limited number of dimensions that can be expected to make most of the difference (**readers interested in an extensive coverage of all possible study factors, as well as an overview of the distribution of these factors across the Carlino-Mills literature are referred to Hoogstra 2005*). Herewith, several potentially important dimensions, such as type of data (cross-section versus panel), functional form, and estimation technique do not display sufficient inter-study or intra-study variance to be considered for investigation in the meta-regression analysis. In addition to the study characteristics that show little variation from study to study and thus have limited value for anything but descriptive purposes, several descriptors of the Carlino-Mills studies exhibit strong interrelationships, thereby holding back the prospect of statistical inference. For instance, the type of region (rural versus urban) studied and spatial detail of the data are strongly correlated with model specification (with the spatial econometric Boarnet model typically being used to examine intra-urban small-area location patterns). Likewise, the robustness of the study results against alternative spatial weights matrices (an issue raised in Boarnet *et al*. 2002 and Boarnet and Chalermpong 2002) can only be assessed from the subgroup of spatial econometric studies, which reduces the sample of observation considerably.

With regard to the group of substantive factors being related to the spatial characteristics of the data, we make a rough distinction between U.S. studies and non-U.S. studies (as with other study dimensions, the available database for the meta-analysis does not allow us to make a more detailed distinction). Furthermore, we distinguish between studies of population and employment location patterns at three levels of spatial detail, namely the extremely large-sized U.S. states, large-sized U.S. BEA regions, and the remaining heterogeneous group of medium and small-sized area units (which is not further disaggregated due to the strong correlation with the different model specifications, see above). With regard to the variable that depicts the temporal setting of the data, a straightforward distinction is made between the study results that refer to the 1960s and 1970s, 1980s, and 1990s. Herewith, the study results by Kristensen and Henry (1997), Holmberg *et al*. (2001), Vergolino and Jatobá (2001), Carruthers and Vias (2003), and Edmiston (2003), which correspond to parts of the 1980s as well as the 1990s are mostly concerned with the 1980s and are classified as such. Finally, a simple dichotomy is adopted that separates between study results based on aggregate (total) population and/ or employment data and data for employment and/or population subgroups (irrespective of the particulars of the groups under consideration).

With respect to the group of methodological study factors we distinguish between observations for which at least one of the standard statistical assumptions of uncorrelated and homescedastic error terms is satisfied and the remaining observations. Next, a distinction is made between studies using unstandardised and standardised population and employment variables, with the latter controlling for size-differences between the spatial units of observations. Two different features related to model specification are discerned. The first is the all-important difference with regard to the specification of left-hand side (LHS) and right-hand side (RHS) population and employment *changes*, *levels*, or the mixture of LHS *changes* and RHS *levels*, which together with above-mentioned issue of variables measurement show the greatest inter-study variance (see final columns of Table 4). As discussed in Section 3.1, the specification reflecting LHS and RHS *changes* mainly reflects the spatial econometric Boarnet model, in which the RHS variable is spatially weighted to control for spill-over (commuting) effects. As stated above, the application of this model also tends to coincide with the use of small-area data for urban regions, in particular, implying that the variation in study results attributed to model specification may actually reflect different sources. In the same vein, eventual deviant study results relating to the model specification of LHS *changes* and RHS *levels* extensively used by Mulligan *et al*. (1999) may also be caused by the focus on 1-year time lags and/ or the lack of exogenous variables. The second feature related to model specification is the number of endogenous RHS variable. Specifically, a broad distinction is made between a regular two-equations system with one RHS endogenous variable in each equation, and more extended frameworks that include additional RHS endogenous variables, reflecting spatial effects, group effects, or previously assumed exogenous factors such as income (Mills and Lubuele 1995), taxes (Danielson and Wolpert 1991), state parks (Duffy-Deno 1997b), endangered species preservation (Duffy-Deno 1997b), and new firms (Edmiston 2003).

Finally, we discriminate between the publication status of two groups of study results, namely those extracted from peer-reviewed journal articles and those that have not (yet) been published in any journal, referring to working papers, book chapters, and the like.

The study features outlined above all provide plausible explanations for the variation in research findings for the jobs-people direction of causality. For instance, several reasons can be put forward to presume that the nature of causality is different between U.S. and non-U.S. (predominantly European) studies. Among these reasons is the greater flexibility of U.S. labour markets where, for instance, the response to changes in labour demand seems to take mainly place via migration, instead of changes in labour participation (see, e.g., Blanchard and Katz 1992; Decressin and Fatas 1995; Broersma and Van Dijk 2002). The substantive interest in the size of the spatial units under examination relates to the spatial scale at which population-employment interaction operates. The area units, whether census tracts, municipalities, or states are essentially arbitrary groupings, and the data within can be aggregated in an infinite number of ways. A problem that arises from the imposition of artificial units of spatial reporting on any continuous geographical phenomenon is the generation of artificial spatial patterns, which is known as the 'modifiable area unit problem' (Heywood *et al*. 1998). The practical implication is that alternative aggregations of the data probably lead to different results, especially so since the strength of the population-employment relationship is known to change over distance (see, e.g., Wheeler 2001). Similar to the geographical characteristics of the data shaping the research outcomes, it seems highly feasible to suggest that the relationship alters over time, due to changing preferences for industrial and residential location, and changing economic conditions to act upon these preferences. For instance, it has been assumed that the 'people follow jobs' direction of causality can be mainly associated with the traditional 'industrial society', whereas the current 'knowledge society' is usually being linked with 'jobs follow people' (see, e.g., Vias 1999; Holmberg *et al*. 2001; Florida 2002). In a similar vein, question marks have already been raised with respect to the robustness of the research findings against the examination of different population and/or employment groups. It can be maintained that the use of 'aggregate' population and employment data yields an average estimate for direction of causality that may conceal contrasting research findings for subgroups.

As for the above-mentioned methodological study features, several display intra-study variance in addition to the usual inter-study variance, thereby clearly indicating that researchers have already explicitly drawn upon the sensitivity of their research findings against alternative study methods. Notably, the effect of using levels as compared to densities as measures of the population and employment variables has attained considerable attention in the literature (see Bao 1996; Glavac *et al*. 1998; Mulligan *et al.* 1999), and for which the findings generally seem to point towards considerable changes in research outcomes. The absence of additional endogenous variables referring to, e.g., spatial effects in many of the model specifications may result in an omitted variable bias, with side-effects on the accuracy of the estimation results that reveal the jobs-people direction of causality. Similarly, it can be assumed beforehand that the estimation results of the models for which the accompanying disturbance terms satisfy the standard statistical criteria of being uncorrelated and homoscedastic are more precise.

The potential impact of publication status on study results is a reiterated issue in studies resorting to meta-analysis for research synthesis. Although publication of a study does not itself affect the research outcome, it may reflect the selection criteria and reporting proclivities of the authors, reviewers and editors who decide if and how a study will be published (Cooper and Hedges 1994). Specifically, researchers may have a tendency to self-censor the publication of 'negative' or statistically insignificant results, a practice that may be invigorated by editorial selection processes (Florax 2002). It has already been revealed for research areas like marketing (see Hubbard and Armstrong 1992), psychology, and medicine (see Sterling *et al*. 1995) that peer review is biased against the publication of such results.

4. Meta-regression analysis

4.1 Set-up

Following the separate discussions of the study results and the associated study characteristics, this section proceeds with the examination of their relationships in order to pass judgement on the empirical evidence for the jobs-people direction of causality. Multivariate regression techniques are employed to evaluate the impact of each of the selected substantive, methodological, and extrinsic study factors separately while controlling for the possible influences of all the other factors. By doing so, rigorous insights can be obtained as to which study features cause most of the variation in research findings that indicate the jobs-people direction of causality (preliminary insights in the pairwise relationships between the study results and underlying study characteristics can be obtained from Table 4). Because the study results refer to four discrete categories, the multivariate analyses are performed using a multinomial logistic regression model, which is well suited to predict the likelihood of a categorical outcome variable given a number of explanatory variables. This model comprises three equations in which the respective dependent variables are defined as the log odds that the estimation results indicate either ‘no interaction’, ‘people follow jobs’, or ‘dual causality’, using ‘jobs follow people’ as the reference category. From each group of study factors that serve as explanatory variables one category is omitted to compare against. The estimated regression coefficients reveal the additive effect of each category compared to 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 change in odds, which can be estimated by exponenting these coefficients (i.e., taking the antilog with the base e). A positive coefficient means that the factor is greater than 1, thereby revealing an increase in the odds. A negative coefficient, on the other hand, complies with a factor that is less than 1, which means that the odds are decreased. In case the coefficient is not significantly different from zero the factor equals 1, which leaves the odds unchanged. The multivariate logistic regression model is estimated using both the unweighted and weighted study samples.[6] The latter serves to warrant that the findings of the meta-analysis are not particularly biased due to putting too much weight on one of the selected studies that is disproportionately represented. It also helps to alleviate a potential bias due to lacking independence among the study results. That is, the sampled results from one and the same study (or cluster of studies) may display a systematic relationship, because they typically

[6] Because its threesome distinction is no longer meaningful after weighting the study results (with the results for the U.S. states and BEA-regions only coming from Mulligan *et al*. 1999), the variable depicting spatial unit size will only be included in the unweighted analyses.

reflect the same data, model specification, and most other underlying study characteristics. But also between different studies (or clusters of studies) there are common characteristics with regard to, e.g., the space-time coverage, research design, and estimation procedures that may cause dependence among seemingly independent study results. According to Florax (2002) the latter form of dependence (also referred to as 'between-study dependence') is usually sufficiently accounted for by means of variability in the study characteristics that specify the heterogeneity of studies. In contrast, the form of dependence generally referred to as 'within-study dependence' is typically more problematic, as it may cause inferences about the significance of the effects to be inaccurate.[7]

Table 4 *Distribution of study results across selected study features*

	(a)				(b)				(a)	(b)
	I	**II**	**III**	**IV**	**I**	**II**	**III**	**IV**	*n*	*n*
substantive study factors										
non-US	21.1	28.9	23.7	26.3	7.2	49.3	10.1	33.3	38	70.0
US *	26.3	31.9	28.1	13.7	26.1	44.1	11.8	18.1	270	238.0
states (US)	28.0	18.0	52.0	2.0					50	
BEA regions (U.S.)	18.0	24.0	58.0	0.0					50	
other *	26.9	36.5	14.4	22.1					208	
1960s & 1970s	22.2	38.9	25.0	13.9	10.2	61.2	10.2	18.4	72	48.6
1990s	27.1	29.2	25.0	18.8	34.1	39.0	9.8	17.1	48	41.4
1980s *	26.6	29.3	29.3	14.9	22.0	43.1	11.9	22.9	188	217.9
groups	37.5	35.4	14.6	12.5	32.1	35.8	17.0	15.1	48	53.4
non-groups *	23.5	30.8	30.0	15.8	19.7	47.2	10.2	22.8	260	254.5
methodological study factors										
corrected errors	30.5	40.7	15.3	13.6	21.8	43.6	8.9	25.7	59	100.4
uncorrected errors *	24.5	29.3	30.5	15.7	22.1	46.2	12.0	19.7	249	207.5
standardised	24.1	31.7	22.8	21.4	11.9	38.5	19.3	30.4	145	135.8
unstandardised *	27.0	31.3	31.9	9.8	29.7	51.2	4.7	14.5	163	172.2
levels-levels	10.7	58.9	10.7	19.6	5.2	54.5	9.7	30.6	56	133.8
changes-levels	19.2	25.7	40.1	15.0	19.5	46.3	15.9	18.3	167	82.6
changes-changes *	48.2	24.7	14.1	12.9	47.8	31.5	8.7	12.0	85	91.5
endogenous 2+	35.6	33.9	16.9	13.6	29.3	40.4	8.1	22.2	59	100.1
endogenous 1 *	23.3	30.9	30.1	15.7	18.4	47.8	12.6	21.3	249	207.8
extrinsic study factors										
non-journal article	38.1	33.9	8.5	19.5	25.6	47.7	4.5	22.2	118	175.7
journal article *	17.9	30.0	39.5	12.6	17.4	42.4	19.7	20.5	190	132.3
overall	25.6	31.5	27.6	15.3	21.8	45.5	11.4	21.4	308	308

I = no interaction; II = jobs follow people; III = people follow jobs; IV = dual causality

[7] Future work may involve the application of formal statistical tests for the presence of dependence among study results, which is thus far practically non-existent in meta-analyses (Florax 2002). Herewith, the multidimensional character of autocorrelation in a meta-analysis implies that the same tests may be applied as for the detection of correlation between observations in the spatial domain. Complications arise from the use of categorical data, with the development of estimation procedures for logistic regressions with spatial dependence still being in its infancy stage (see for applications, e.g., Dubin 1997; Lin 2003).

See Table 3 for the meaning of study samples (a) and (b). Light and dark shading indicate cell proportions being more than 10.0 below and above, respectively, the corresponding overall share. * Reference categories in the multivariate regression model.

Table 5 *Estimation results multivariate analyses*

dependent variable:		$\log\left(\frac{Prob\ (\mathrm{I})}{Prob\ (\mathrm{II})}\right)$			$\log\left(\frac{Prob\ (\mathrm{III})}{Prob\ (\mathrm{II})}\right)$			$\log\left(\frac{Prob\ (\mathrm{IV})}{Prob\ (\mathrm{II})}\right)$		
		B	Exp(B)		B	Exp(B)		B	Exp(B)	
intercept	(a)	0.335			-0.094			-1.090		
	(b)	1.247		**	-0.484			-2.551		***
substantive study factors										
non-U.S.	(a)	-1.447	0.235	**	1.822	6.183	***	0.482	1.620	
	(b)	-3.139	0.043	***	-0.499	0.607		0.838	2.311	*
states	(a)	1.401	4.058	**	1.832	6.243	***	-2.322	0.098	*
	(b)									
BEA regions	(a)	0.619	1.858		1.714	5.550	***	-21.859	0.000	***
	(b)									
1960s/1970s	(a)	0.053	1.054		-1.487	0.226	***	-0.154	0.857	
	(b)	0.211	1.234		-0.019	0.981		0.020	1.020	
1990s	(a)	0.739	2.094		-0.448	0.639		0.468	1.596	
	(b)	3.722	41.350	***	2.079	7.995	**	0.477	1.611	
groups	(a)	0.722	2.058		0.606	1.834		-1.065	0.345	
	(b)	1.291	3.635	**	-0.225	0.799		-1.164	0.312	*
methodological study factors										
correct. errors	(a)	-0.765	0.466		-0.553	0.575		-0.938	0.392	*
	(b)	-1.339	0.262	***	-1.166	0.312	**	-0.008	0.992	
standardised	(a)	1.058	2.881	**	-0.975	0.377	**	1.164	3.203	**
	(b)	1.604	4.972	**	2.489	12.048	***	1.602	4.964	***
levels-levels	(a)	-3.540	0.029	***	-0.524	0.592		-1.013	0.363	
	(b)	-5.252	0.005	***	-2.145	0.117	**	0.789	2.222	
changes-levels	(a)	-2.151	0.116	***	0.437	1.548		0.450	1.568	
	(b)	-3.102	0.045	***	-1.263	0.283		0.723	2.060	
2+ endogenous	(a)	-0.754	0.470		0.218	1.243		-0.010	0.990	
	(b)	0.291	1.337		0.702	2.017		1.332	3.788	**
extrinsic study factors										
non-article	(a)	1.102	3.011	**	-1.352	0.259	**	0.717	2.048	
	(b)	-0.554	0.574		-2.135	0.118	***	-0.356	0.700	

I = no interaction; II = jobs follow people; III = people follow jobs; IV = dual causality

See Table 2 for the meaning of study samples (a) and (b) and Table 4 for reference categories.

* < 0.10, ** < 0.05, *** < 0.01

4.2 Results multivariate analyses

The main findings of the multivariate logistic regression analysis are the following (see Table 5). It is revealed that the likelihood of a research finding pointing towards ‘no interaction’ instead of ‘jobs follow people’ is largely affected by the choice of a particular model specification (relating to population and employment levels/ changes), the spatial setting of the data, and variables measurement. Specifically, the application of a model that measures LHS and RHS *levels* or LHS *changes* and RHS *levels* seems less likely (albeit the latter specification to a lesser extent) to produce research findings that reveal the lacking of any form of population-employment interaction compared to a model specification that measures both LHS and RHS *changes* (baseline category). As for the geographical characteristics of the data, the negative coefficients associated with the use of non-U.S. data indicate that the log odds to find 'no interaction' instead of 'jobs follow people' decrease significantly when the region under examination is situated outside the United States. Following the estimated coefficient in the weighted study sample, such an examination is about 23 times less likely (=1/0.043) to fail in detecting any sort of interaction as opposed to a U.S.-oriented study. Similarly, using highly aggregate spatial data at the level of U.S. states appears to be some 4 times more likely to yield findings indicative of 'no interaction' compared to data at a finer spatial level (including BEA regions). As far as the remaining study factors are concerned there is some evidence, albeit less robust, that the publication status, data characteristics, and appropriateness of statistical quality of the modelling affect the likelihood to come across 'no interaction' instead of 'jobs follow people'. After weighting the observations in the study sample, the estimated coefficient associated with a 'non-journal article', which otherwise seems to indicate that such a publication outlet is more likely than a journal article (baseline category) to reveal 'no interaction', is no longer statistically significant. In a similar vein, the findings that suggest that the odds to find 'no interaction' are increased significantly by using a particular dataset (referring to the 1990s and reflecting distinct population and/ or employment groups) and decreased significantly by making corrections for heteroscedastic and/ or autocorrelated error terms are only observable in the weighted study sample. Finally, no statistical evidence is found in both the unweighted and weighted study samples that the log odds are changed as a result of adopting an extended model specification (with additional RHS endogenous variables) instead of a regular specification.

From Table 5 it appears that the likelihood to find to 'people follow jobs' instead of 'jobs follow people' increases when large-scale data at the level of U.S. states or BEA regions instead of more detailed spatial data are examined, which explains to a large extent why the study by Mulligan *et al.* (1999) shows such a considerable bias. There is also evidence that the use of unstandarised variables has contributed significantly to the deviance in study results by Mulligan *et al.* (1999), which follows from the negative significant coefficient for the standardised variables, which indicates that the odds decrease when such variables are employed, and consequently increase when unstandardised variables (baseline category) are employed. The coefficient for the standardised variables switches in sign when the abovementioned study and other studies in the sample are weighted, thereby indicating that the effect of 'variables measurement' is not unambiguous, but very much dependent on the study context. Next, it is shown that log odds are significantly different between the diverse publication outlets, with the estimated regression coefficients suggesting that peer review is prejudiced in favour of the publication of study results indicative of 'people follow jobs'. As far as the remaining study factors are concerned, there are signs that the presence of heteroscedasticity and autocorrelation among the error terms, the spatial setting of the data, and the time period covered by the data may affect the odds to find evidence for one-way causality running from employment to population. Remarkably, it follows from the estimations based on the unweighted study sample that the use of data that refer to the 1960s and 1970s, instead of the 1980s (baseline category) increases such odds significantly. Together with the finding of a negative coefficient being associated with the 1990s in the weighted study sample, the estimation results run counter to the widely accepted supposition that one-way causality running from population to employment (i.e., people follow jobs) have become more important over time. With the effects of spatial unit selection, variables measurements, and temporal setting of the data being controlled for there is no evidence found that the particular model specification of LHS *changes* and RHS *levels* being adopted by Mulligan *et al*. (1999), among others, assert any influence on the odds to find 'people follow jobs' instead of the other way around. Similarly, no presence of group effects or influences being attributable to the number of endogenous variables included in the specification can be observed.

The regression coefficients of the final equation in Table 5 indicate that the chances to find evidence for 'dual causality' are significantly reduced when large-sized units like U.S. states and BEA regions are selected as spatial units of observation, which is hardly surprising given that such a combination is extremely uncommon (see Table 4). The estimated coefficients also reveal that standardising the variables greatly enhance the likelihood that the model parameters indicate two-way interaction instead of 'jobs follow people'. As for the coefficients that indicate whether corrections for heteroscedastic and/ or autocorrelated error terms make any difference, there is some evidence that such corrections lessen the chance to find 'dual causality'. A statistically significant effect can only be observed, though, in the unweighted study sample. Likewise, proof that the spatial setting of the data, the population and employment characteristics of the data, and the number of endogenous variables included in the model are affecting the odds is only given in the weighted sample. Finally, irrespective whether the observations in the study sample are weighted or not, the variables depicting publication status, model specification (levels/ changes), and time period covered by the data do not show a statistically significant relationship with the dependent variable.

5. Conclusions

This study has revealed that the literature on population-employment interaction show considerable variation in research findings indicating the direction of causality, which confirms popular belief. The findings provide above all support for the 'jobs follow people' hypothesis, notably when the numerous study results coming from one and the same study and cluster of associated studies are controlled for. Widespread support for the hypothesis of employment driving population changes ('people follow jobs') is markedly absent, with most of the evidence being based on a single study by Mulligan *et al*. (1999). Results from the meta-regression analysis have indicated that from the list of substantive study factors especially the spatial detail of analysis plays a crucial role in shaping the research findings that indicate the jobs-people direction of causality. Specifically, it has been shown that the chances to find 'people follow jobs' and 'dual causality' significantly increase and decrease, respectively, when large-scale data at the level of U.S. states and BEA regions are used. Next, there is some evidence that the study results coming from U.S. and non-U.S. oriented studies are significantly different from each other, thereby firmly hinting at social, cultural, and institutional differences shaping the nature of population-employment interaction. Evidence that the examination of different time periods or the examination of data referring to specific employment and/ or population groups make much of a difference has been noticeably absent, albeit that this may have everything to do with the rather crude categorisation being applied. As for the different methodological study factors, it has been revealed that the different model specifications being employed in the Carlino-Mills literature are largely responsible for the variation in study results. Notably, the finding of ‘no interaction’ appears to be associated with the model speciation in which the essential population and employment variables are measured as changes instead of levels. As the application of this model tends to coincide with the examination of growth patterns (mainly urban) at an extremely fine spatial detail, it remains undecided what the exact reasons are for the strong bias towards the research finding of ‘no interaction’. It appears, however, that the modelling of population and employment dynamics is more difficult than that of static levels, and in combination with the fine spatial scale, which requires the tricky task to control for spillover effects between locations, especially prone to arrive at statistically insignificant estimates suggesting the absence of interaction. The use of standardised versus unstandardised variables is another methodological feature for which a systematic relationship with the study results indicative of population-employment interaction has been revealed. Herewith, the relationship appeared to change as different study samples were investigated, thereby giving rise to the supposition that the effect of standardising the population and employment numbers by the area size of the spatial units depends on other study characteristics such as the actual size of these spatial units and the type of region (high-density versus low-density) under examination. The meta-analyses generally did not reveal any substantiation for the suppositions that by failing to satisfy the criteria of independent and identically distributed error terms or by applying an extended model specification with additional endogenous variables significantly alters the study results. Similarly, by holding the other study characteristics constant no evidence was found for any sort of publication bias in the study results.

The following suggestions can be made as for the issues that warrant further investigation. First, the effects of many of the study factors discerned did not seem to be independent from each other. For instance, it has been speculated above that the effect of standardising population and employment levels by area measures may work out differently for studies concentrating on different type of regions and with varying degrees of spatial detail. Similarly, it seems reasonable to assume that, for instance, the effects of the time period and region type under examination are conditional upon each other. Given the limited data available, the investigation of such interaction effects was not possible in this study, but may be pursued in future meta-analytical work. Second, in addition to the incorporation of interaction effects as mentioned above, the explanatory power of the meta-regression model may greatly improve by the inclusion of other features displaying intra-study or inter-study variance. Herewith, future research may especially take advantage from recent advances with regard to the functional form being assumed (non-linear instead of linear) and the type of data being used (panel data instead of cross-sectional data), which reflect potentially important study factors. Similarly, future work may benefit from an increase in studies in order to assess the effects of substantive study characteristics like the time period, region type, and employment type under examination more rigorously. As stated above, the rather broad categorisation with regard to the time period under examination and the simple distinction between using aggregate data and data that refer to distinct employment and/ or population groups (irrespective of the sort of groups), for instance, leave much to be desired and can certainly be held responsible for some of the remaining unexplained variation in study results.

In sum, while this study has taken a massive step forward by sifting through potential substantive, methodological, and extrinsic study factors influencing the study results indicative of population-employment interaction to identify those that are significant, it also highlights the need for stronger future research designs. Somewhat ironically, the opportunities available to meta-analysis to make sense of the rapidly expanding literature on population-employment interaction are enhanced by a continuing increase of these studies. Alternatively, the data required for a meta-analysis can also be self-generated, which is shown in a supplementary paper (see Hoogstra *et al*. 2004).

References

- Arauzo-Carod, J.M. (2002), *Determinants of population and jobs at a local level. An application for Catalan municipalities*. Working paper.

- Argo, L.R. (2001), *Challenges for the New West. Economic impacts of wilderness on Nevada's rural counties*. Master's thesis, University of Nevada.

- Bao, S. (1996), *Restructuring in functional economic areas: A GIS based study of rural linkages to urban areas in selected Southern regions*. Ph.D. dissertation, Clemson University.

- Bao, S., M.S. Henry, and D.L. Barkley (1999), *Identifying urban-rural linkages. Tests for spatial effects in the Carlino-Mills model*. Working paper.

- Barkley, D.L., M.S. Henry, and S. Bao (1998), The role of local school quality in rural employment and population growth. *The Review of Regional Studies*, 28, pp. 81-102.

- Bijmolt, T.H.A. and R.G.M. Pieters (20010, Meta-analysis in marketing when studies contain multiple measurements. *Marketing Letters*, 12, pp. 157-169.

- Blanchard, J.B. and L. Katz (1992), Regional Evolutions. *Brookings Papers on Economic Activity*, 1, pp. 1-61.

- 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. (1994a), An empirical model of intrametropolitan population and employment growth. *Papers in Regional Science*, 73, pp. 135-152.

- Boarnet, M.G. (1994b), The monocentric model and employment location. *Journal of Urban Economics*, 36, pp. 79-97.

- Boarnet, M.G. and S. Chalermpong (2002), *Urban growth and highway infrastructure: A "before" and "after" test of the influence of highways on metropolitan growth patterns*. Working paper.

- Boarnet, M.G., S. Chalermpong, and E. Geho (2002), *Specification issues in models of population and employment growth*. Working paper.

- 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, pp. 179-204.

- Bollinger, C.R. and K.R. Ihlanfeldt (2001), *Spatial interactions among employment and population groups*. Working paper.

- Broersma, L. and J. van Dijk (2002), Regional labour market dynamics in the Netherlands. *Papers in Regional Science*, 81, pp. 343-364.

- Carlino, G.A. and E.S. Mills (1987), The determinants of county growth. *Journal of Regional Science*, 27, pp. 39-54.

- Carruthers, J.I. and A.C. Vias (2003), *Urban, suburban, and exurban sprawl in the Rocky Mountain West*. Working paper.

- Carruthers, J.I. and G.F. Mulligan (2004), *Land absorption in United States metropolitan areas: Estimates and projections from regional adjustment models*. Working paper.

- Clark, D.E. and C.A. Murphy (1996), Countywide employment and population growth: An analysis of the 1980s. *Journal of Regional Science*, 36, pp. 235-256.

- Cooper, H. and L.V. Hedges (1994), *The handbook of research synthesis*. New York: Sage Publishers

- Dalhuisen, J.M., R.J.G.M. Florax, H.L.F. de Groot, and P. Nijkamp (2003), Price and income elasticities of residential water demand: A meta-analysis. *Land Economics*, 79, pp. 292-308.

- Danielson, M.N. and J. Wolpert (1991), Distributing the benefits of regional economic development. *Urban Studies*, 28, pp. 393-412.

- Decressin, J. and A. Fatas (1995), Regional labor market dynamics in Europe. *European Economic Review*, 39, pp. 1627-1655.

- Deitz, R. (1998), A joint model of residential and employment location in urban areas. *Journal of Urban Economics*, 44, pp. 197-215.

- Deller, S.C., T-H. Tsai, and D.W. Marcouiller (2001), The role of amenities and quality of life in rural economic growth. *American Journal of Agricultural Economics,* 83, pp. 352-365.

- Dubin, R.A. (1997), A note on the estimation of spatial logit models. *Geographical Systems*, 4, pp. 181-194

- Duffy-Deno, K.T. (1997a), Economic effect of endangered species preservation in the non-metropolitan West. *Growth and Change*, 28, pp. 263-288.

- Duffy-Deno, K.T. (1997b), The effect of state parks on the county economies of the West. *Journal of Leisure Research*, 29, pp. 201-224.

- Duffy-Deno, K.T. (1998), The effect of federal wilderness on county growth in the intermountain Western United States. *Journal of Regional Science*, 38, pp. 109-136.

- Edmiston, K.D. (2003), *The net effects of large plant locations and expansions on county employment*. Working paper.

- Florax, R.J.G.M. (2002), *Accounting for dependence among study results in meta-analysis: methodology and applications to the valuation and use of natural resources*. Free University of Amsterdam research memoranda 2002-5.

- Florax, R.J.G.M., H.L.F. de Groot, and R.A. de Mooij (2002), *Meta-analysis in policy-oriented economic research*. CPB report.

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

- Glass, G.V. (1976), Primary, secondary, and meta-analysis of research. *Educational Research*, 5, pp. 3-8.

- 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, pp. 632-656.

- Hedges, L.V. (1997), The promise of replication in labour economics. *Labour Economics*, 4, pp. 111-114.

- Hedges, L.V. and I. Olkin (1980), Vote-counting methods in research synthesis. *Psychological Bulletin*, 88, pp. 359-369.

- Hedges, L.V. and I. Olkin (1985), *Statistical methods for meta-analysis*. New York: Academic Press.

- Henry, M.S., K. Kristensen, D.L. Barkley, and S. Bao (1999), Extending Carlino-Mills models to examine urban size and growth impacts on proximate rural areas. *Growth and Change*, 30, pp. 526-548.

- Henry, M.S., D.L. Barkley, and S. Bao (1997), The hinterland's stake in metropolitan growth. Evidence from selected southern regions. *Journal of Regional Science*, 37, pp. 479-501.

- 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, pp. 171-193.

- Heywood, I., S. Cornelius, and S. Carver (1998), *Introduction to geographical information systems*. New York: Addison Wesley Longman.

- Holmberg, I. and U. Strömquist (1998), *LOSE - Länstrafik och samhällsekonomi* (Public Transport in Counties and Economic Consequences). Stockholm: Transportforsknings-beredningen.

- Holmberg, I., B. Johansson, and U. Strömquist (2001), *A simultaneous model of long-term regional job and population changes*. Working paper.

- Hoogstra, G.J. (2005), *Spatial dynamics of population-employment interaction. An analysis of local growth patterns in the Northern Netherlands*. Ph.D. dissertation, University of Groningen (forthcoming).

- Hoogstra, G.J., R.J.G.M. Florax, and J. van Dijk (2004), *Specification issues and sample selection in population-employment interaction models. A quasi-experimental meta-analysis*. Working paper.
- Hubbard, R., and J.S. Armstrong (1992), Are null results becoming an endangered species in marketing? *Marketing Letters*, 3, pp. 127-136.
- Kristensen, K. and M.S. Henry (1997), *Peripheral population and employment change in Denmark's functional economic regions: Does the hinterland have a stake in urban growth*? Working paper.
- Lin, G. (2003), A spatial logit association model for cluster detection. *Geographical Analysis*, 35, pp. 329-340.
- Luce, T.F. (1994), Local taxes, public services, and the intrametropolitan location of firms and households. *Public Finance Quarterly*, 22, pp. 139-167.
- McCloskey, D.N. (1985), The loss function has been mislaid: The rhetoric of significance tests. *American Economic Review*, 75, pp. 201-205.
- McDonald, J.F. (1992), Assessing the development status of metropolitan areas. In: E.S. Mills and J.F. McDonald (eds.), *Sources of metropolitan growth*. New Brunswick, NJ: Center for Urban Policy Research.
- Mills, E.S. and G.A. Carlino (1989), Dynamics of county growth. In: A.E. Andersson *et al*. (eds.), *Advances in spatial theory and dynamics*. Studies in Regional Science and Urban Economics, vol. 20. Amsterdam: Elsevier Science.
- Mills, E.S. and L.S. Lubuele (1995), Projecting growth of metropolitan areas. *Journal of Urban Economics*, 37, pp. 344-360.
- Mulligan, G.F., A.C. Vias, and S.M. Glavac (1999), Initial diagnostics of a regional adjustment model. *Environment and Planning A*, 31, pp. 855-876.
- Nzaku, K. and J.O. Bukenya (2004), *The influence of amenities on economic development in Alabama*. Working paper.
- Pantuosco, L., D. Parker, and G. Stone (2001), The effect of unions on labor markets and economic growth. An analysis of state data. *Journal of Labor Research*, 22, pp. 195-205.
- Rosenberger, R.S., T.G. Gebremedhin, and Y. Hailu (2002), *An economic analysis of urbanization of agricultural land in West Virginia*. Working paper.
- Rudner, L., G.V. Glass, D.L. Evartt, and P.J. Emery (2002), *A user's guide to the meta-analysis of research studies*. ERIC Clearinghouse on Assessment and Evaluation, University of Maryland, College Park.
- Schmitt, B. (2000), Economic geography and contemporary rural dynamics: An empirical test on some French regions. *Regional Studies*, 33, pp. 697-711.
- Schmitt, B. and M.S. Henry (2000), Size and growth of urban centers in French labor market areas. Consequences for rural population and employment. *Regional Science and Urban Economics*, 30, pp. 1-21.

- Schmitt, B., M.S. Henry, M. Hilal, and V. Piguet (2000), *Urban spatial externalities. Do they affect rural areas*? Working paper.

- Schmitt, B., M.S. Henry, and V. Piguet (1999), *Spatial econometric models for simultaneous systems: Comparison of some models for selected regions in France*. Working paper.

- Schmitt, B., M.S. Henry, and V. Piguet (2002), *Urban growth impacts on rural population and employment in export and service industries in functional economic regions*. Working paper.

- Sohn, J. and G.J.D. Hewings (2000), *Spatial evidence of agglomeration economies in Chicago*. Working paper.

- Stanley, T.D. (2001), Wheat from chaff: Meta-analysis as quantitative literature review. *Journal of Economic Perspectives*, 15, pp. 131-150.

- Stanley, H. and S. Jarrell (1989), Meta-regression analysis. A quantitative method of literature surveys. *Journal of Economic Surveys*, 3, pp. 161-170.

- Sterling, T. D., W.L. Rosenbaum, and J.J. Weinkam (1995), Publication decisions revisited: The effect of the outcome of statistical tests on the decision to publish and vice versa. *The American Statistician*, 49, pp. 108-112

- Suriñach, J. J.C. Duque, R. Ramos, and V. Royuela (2003), Publication patterns in regional and urban analysis: Have topics, techniques and applications changed during the 1990s. *Regional Studies*, 37, pp. 351-363.

- Summers, J.O. and D.B. MacKay (1976), On the validity and reliability of direct similarity judgments. *Journal of Marketing Research*, 13, pp. 289-295.

- Van den Bergh, J.C.J.M., K.J. Button, P. Nijkamp, and G.C. Pepping (1997), *Meta-analysis in environmental economics*. Dordrecht: Kluwer Academic Publishers.

- Vergolino, J.R. and A.D.S.C. Jatobá (2001), *Os determinantes do crescimento da população e do emprego das microrregiões do Nordeste do Brasil – 1970/1996. Uma abordagem econométrica*. Working paper.

- Vias, A.C. (1998), *An analysis of population and employment growth in the nonmetropolitan Rocky Mountain West, 1970-1995*. Ph.D. dissertation, Arizona University.

- Vias, A.C. (1999), Jobs follow people in the Rural Rocky Mountain West. *Rural Development Perspectives*, 14, pp. 14-23.

- Vias, A. C. and G.F. Mulligan (1999), Integrating economic base theory with regional adjustment models. The metropolitan Rocky Mountain West. *Growth and Change*, 30, pp. 507-525.

- Wheeler, C.H. (2001), A note on the spatial correlation structure of county-level growth in the U.S. *Journal of Regional Science*, 41, pp. 433-449.

- Wolf, F.M. (1986), *Meta-Analysis: Quantitative methods for research synthesis*. Thousand Oaks, CA: Sage Publications.

- York, A. and D. Munroe (2004), *An integrated regional analysis of land-use change in Southern Indiana: Extending the Carlino-Mills model*. Working paper.