



Residential relocation in response to light rail transit investment: case study of the Hudson–Bergen Light Rail system

Xiaobo Liu¹ · Yi Deng² · Scott Le Vine^{1,3,4}

Received: 5 November 2015/Revised: 16 February 2016/Accepted: 28 March 2016/Published online: 19 April 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract It is widely acknowledged that the improved accessibility enabled by investment in public transport services can, under favorable market conditions, impact the local real estate market within the zone of influence of the service's stations. The motivation for this study is to establish the nature of two such impacts, specifically the spatial and socio-economic patterns of residential relocations that are driven by the new light rail transit (LRT) service. Using empirical data ($n = 1,023$) from the Hudson–Bergen Light Rail system in New Jersey (US), we report findings regarding the impacts of the introduction of the new LRT service. We investigate two linked dimensions; the first is the distinctive socio-economic profile of LRT passengers who self-report having relocated to the new transit corridor due, at least in part, to the new transit service. The second is their proximity (following their residential relocation) to the new LRT line's stations. We present a novel analysis that accounts for endogeneity between these two dimensions of residential relocation. Of light rail passengers who engaged in a residential relocation in the 5 years prior to the survey, two-thirds (69 %) indicate that proximity to the light rail service was a 'somewhat' or 'very' important consideration. Via the multivariate analysis, we demonstrate that small household size, low income, youth (as opposed to

older age), and low car ownership are each positively linked, *ceteris paribus*, with having engaged in a residential relocation motivated by the new transit service. Finally, higher household income is found to be associated with distance (after relocation) to the nearest transit station, which is consistent with bid-rent theory.

Keywords Light rail transit · Residential relocation · Transit-oriented development

1 Introduction

Light rail transit (LRT) has attracted increasing interest from policymakers in recent decades, with new services being introduced across a range of metropolitan regions. The improved accessibility engendered by LRT is seen by many observers as a catalyst for transit-oriented developments (TODs), and indeed in many cases residential development accompanies the introduction of new LRT services [1, 2]. The general issue encapsulated by the TOD concept is that the stimulus of introducing LRT is thought to have non-trivial impacts on the real estate market of the service area.

The motivation for this study was to investigate the dynamics of residential relocation associated with the inauguration of new LRT service. Empirical data are sourced from a 2008 survey of passengers ($n = 1,023$) of the Hudson–Bergen Light Rail (HBLR) system in New Jersey (US). Fieldwork for the survey was undertaken eight years after service inauguration, which implies that this study's results should be interpreted as mid-term effects after the residential property market has had the chance to initially re-equilibrate in response to the stimulus of the new LRT system. The HBLR system (see Fig. 1) feeds heavy rail and ferry services

✉ Xiaobo Liu
Xiaobo.liu@swjtu.cn

¹ School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 610031, People's Republic of China

² Parsons Transportation Group, 100 Broadway, New York, NY 10005, USA

³ State University of New York at New Paltz, New Paltz, NY 12561, USA

⁴ Imperial College London, South Kensington, London SW7 2AZ, England, UK

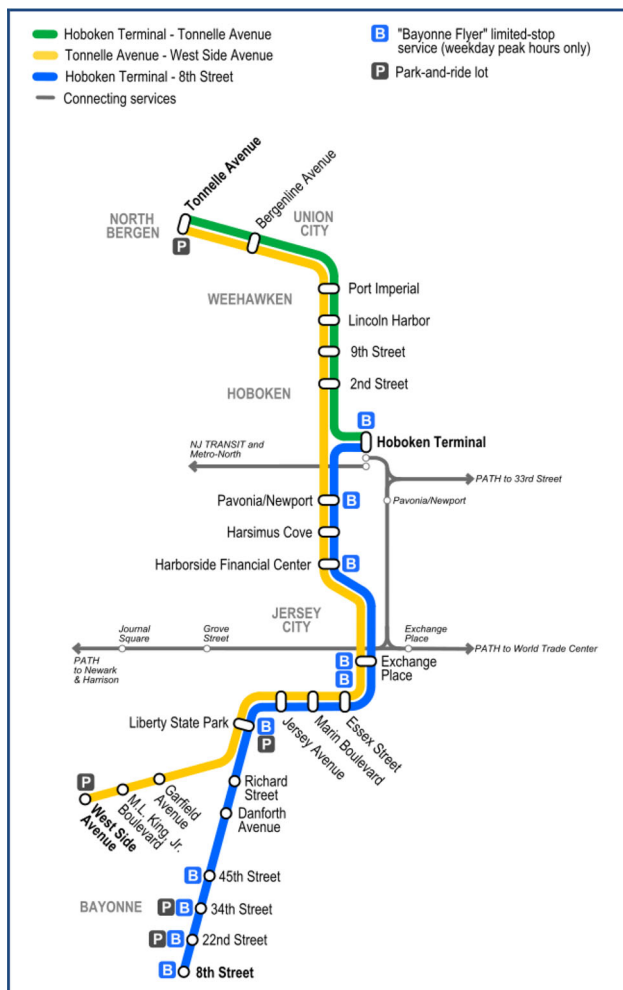


Fig. 1 Schematic representation of the Hudson–Bergen Light Rail (HBLR) system

that connect to midtown and downtown Manhattan, and also provides local connectivity between pre-automobile urban neighborhoods of Hudson County, New Jersey (which is characterized by the sixth-highest residential density among counties in the US). In a study of the land use impacts of the HBLR system, Robins and Wells [3] estimate approximately 10,000 newly built residential units, representing a gross investment of \$5.3 billion (in 2008 prices).

The results reported in this paper are twofold. First, we investigate the distinctive socio-economic profile of LRT passengers who self-report having relocated to the new transit corridor due, at least in part, to the new transit service. Second, we document the spatial patterns (proximity to LRT stations) of residential relocations which were self-reported to have been influenced by the introduction of the LRT system, as a function of socio-economic characteristics. Both descriptive statistics and results from multivariate analyses are presented.

Although the characteristics of transit-oriented development and the impact of transit on property values are well documented (cf. [4]), few studies have explicitly investigated the impact of public transport investment on residential relocation behavior. Table 1 contains a summary listing of the key characteristics of relevant earlier studies. Specifically, the contribution of the present paper is to provide new results regarding two research questions that remain under-researched, while accounting for the possibility of endogeneity between them:

1. *What is the socio-economic profile of people most likely to relocate in response to the stimulus of public transport investment? (a new LRT service, in the case study reported in this paper)*
2. *Conditional on an LRT passenger having relocated in the prior 5 years, how do socio-economic characteristics relate with the proximity of their residential choice to the public transport service?*

The rest of this paper is structured as follows. Section 2 introduces the empirical data employed on this study, and Sect. 3 outlines the analytical framework. Section 4 then summarizes and concludes this paper.

2 Empirical data

The empirical data for this study were collected via an intercept-survey undertaken of passengers waiting on platforms at the seven HBLR stations between the Hoboken Terminal and Tonnelle Avenue stations (see Fig. 1) on 1 May, 2008. The questionnaire instrument contained 39 questions organized around three themes:

- travel, employment, and residential patterns and changes,
- customer satisfaction (not considered for the purposes of the present study), and
- socio-economic and demographic characteristics.

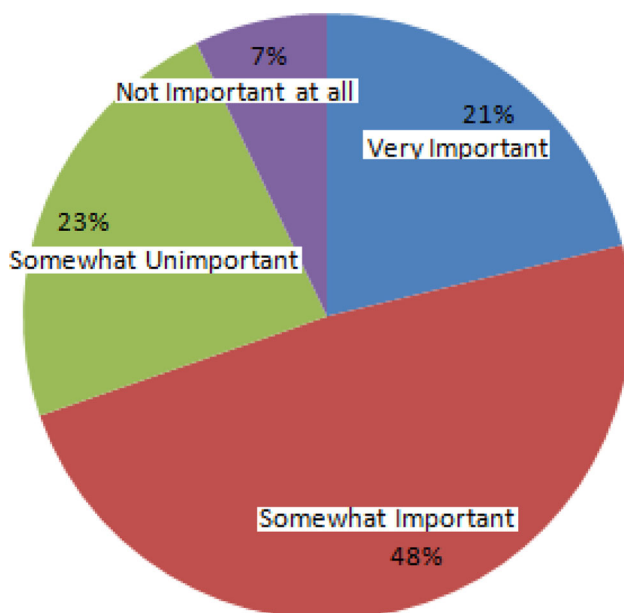
5,384 questionnaires were distributed; the overall response rate was 19 %. Following data processing, a total of 1,023 complete responses for which the respondent's residential address could be successfully geocoded were taken forward for statistical analysis in this study.

Survey respondents reported the length of time that they had lived at their current address, and those who indicated that they lived at their current address for five or fewer years were subsequently asked to indicate how important the availability of the HBLR service was to their residential relocation decision.¹

¹ The question wording was: *How important was the availability of the Hudson–Bergen Light Rail service in your decision to move?*

Table 1 Summary of the previous literature

Reference	Empirical context	Dataset	Key findings
Lund [5]	Los Angeles, San Diego, and San Francisco, California	2003/2004 hand-delivered mail-back survey ($n = 605$)	Among respondents that had recently moved to a Transit-Oriented Development (TOD), access to transit was ranked fifth (out of nine factors) in importance as a factor (after housing quality/type, housing cost, neighborhood quality, and access to shops/services)
Ketraungroch [6]	Bangkok, Thailand	Household Socio-Economic Survey [SES] ($n = 1,445$ households in 1998; $n = 1,512$ households in 2004)	Higher income workers were found to exhibit sharper bid-rent curves than lower income workers, consistent with economic theory
Olaru et al. [7]	Perth, Western Australia (newly opened commuter rail service)	2007 personal-interview survey ($n = 1,034$)	Olaru and colleagues report an estimate that, <i>ceteris paribus</i> , each additional minute of distance from a rail station is associated with a decrease in property value of US \$3,770
Cao and Schoner [8]	Light rail transit in Minneapolis	2011 postal survey ($n = 267$)	Among respondents that had recently moved to the LRT corridor, access to public transport was ranked fourth (of 30 candidate factors) in importance as a factor in residential location choice (after affordability, walkability, and proximity to parks/open space)
Dai and Weinzimmer [9]	San Francisco Bay Area (Silicon Valley)	2013 online survey ($n = 130$) of employer shuttle bus passengers	Among shuttle passengers, proximity to the service was ranked fifth (of 11 candidate factors) in importance as a factor in residential location choice (after Walkability, proximity to cultural amenities, proximity to traditional public transport, and urban-neighborhood amenities)
This study	Hudson–Bergen Light Rail System, New Jersey, US	2008 passenger platform survey ($n = 1,023$)	See Sect. 4

**Fig. 2** Among LRT passengers that report having relocated within the past 5 years, the self-reported ‘importance’ of the LRT service in the relocation decision. Adapted from [10]

Two-fifths (42 %) of respondents reported having relocated within the past 2 years, with a further quarter (24 %) indicating that they relocated more than two but fewer than five years prior. Figure 2 shows the distribution of responses to the importance-of-LRT-in-relocation decision question; 21 % of movers indicated that the LRT service was very important to their relocation decision, and a further 48 % reported that it was somewhat important.

3 Analytical framework

A Type II Tobit discrete–continuous specification [11] was employed in the subsequent quantitative analysis, with the discrete dimension specified to be whether or not a person had relocated to the HBLR corridor within the prior 5 years and the continuous dimension (conditional on a respondent having relocated) the street-network distance of their new residence to the nearest LRT station.

Formally, we denote these dimensions as follows:

$$d_i^* = \alpha'z_i + e_i, \quad i = 1, \dots, N, \quad (1)$$

$$d_i = 1, \quad \text{if } d_i^* > 0; \quad d_i = 0, \quad \text{if } d_i^* \leq 0, \quad (2)$$

$$y_i^* = \beta'x_i + \varepsilon_i, \quad i = 1, \dots, n, \quad n < N, \tag{3}$$

$$y_i = y_i^*, \text{ if } d_i = 1; \quad y_i = 0, \text{ if } d_i = 0, \tag{4}$$

where d_i^* and y_i^* are the latent, continuous dependent variables for individual worker i ; N , and n are the numbers of observations in the full dataset (including respondents who did not relocate) and the subset of respondents who did relocate, respectively. z_i and x_i are the vectors of observed variables that are treated as independent variables in the discrete and continuous models, respectively. Note that the variables within these two vectors are not mutually exclusive; the same variable can appear in both of the modeled dimensions. α' and β' are corresponding vectors of parameters to be estimated, and e_i and ε_i are the disturbance terms which may be correlated through a correlation coefficient ρ that is independent of z_i and x_i . e_i and ε_i are bivariate normally distributed, with zero mean and unknown covariance matrix, denoted in Eq. (5) as follows:

$$\begin{bmatrix} e_i \\ \varepsilon_i \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho\sigma_\varepsilon \\ \rho\sigma_\varepsilon & \sigma_\varepsilon^2 \end{bmatrix} \right), \tag{5}$$

with σ_ε^2 normalized to one for purposes of model identification.

The standard Heckman two-step estimator is employed [11], which is based on the conditional mean expressions and the truncated bivariate normal distribution of the error terms. The expected value of the observed dependent variable y_i is

$$\begin{aligned} \beta'x_i + \rho\sigma_\varepsilon\lambda_i \left(\frac{\alpha'z_i}{\sigma_\varepsilon} \right) E(y_i|x_i, z_i) &= E(y_i^*|x_i, z_i, d_i = 1) \\ &= \beta'x_i + E(\varepsilon_i|e_i > -\alpha'z_i), \end{aligned} \tag{6}$$

which simplifies to

$$\beta'x_i + \rho\sigma_\varepsilon \frac{\phi(\alpha'z_i)}{\Phi(\alpha'z_i)}, \tag{7}$$

where σ_ε is fixed at one; $\lambda_i(\alpha'z_i) = \phi(\alpha'z_i)/\Phi(\alpha'z_i)$ is the inverse Mills ratio; and ϕ and Φ are standard normal probability distribution function and cumulative distribution function, respectively. Equation (7) implies that ignoring the term $\rho\sigma_\varepsilon\lambda_i$ would in effect omit a variable from Eq. (3) under censoring. Thus, the ordinary least squares (OLS) regression of Eq. (2) will yield unbiased estimates of β only if $\rho = 0$ or if correlation between λ_i and x_i is zero. Equation (7) also demonstrates that we could estimate β consistently using the “relocators-only” subset of the survey data by an OLS regression of y_i on x_i and $\lambda_i(\alpha'z_i)$, if α were known). Based on this observation, Heckman’s two-step estimator is calculated by the following procedure:

1. First, the full sample is employed to estimate a binary probit model using standard maximum-likelihood

- techniques, to obtain estimates of α , i.e., $Pr(d_i = 1) = \Phi(\alpha'z_i)$ and $Pr(d_i = 0) = 1 - \Phi(\alpha'z_i)$.
2. Next, $\lambda_i(\alpha'z_i)$ is estimated for each survey respondent that reported having relocated i .

Finally, the sub-sample of only respondents that relocated is used to estimate β and $\beta_\lambda = \rho\sigma_\varepsilon$, by OLS of regressing y_i on x_i and the estimated λ_i .²

The OLS standard error estimates calculated by this estimation procedure require correction, as the error term in Eq. (7) may be heteroskedastic and we use fitted rather than actual values of λ_i . Furthermore, the resulting estimates are consistent, but not asymptotically efficient (i.e., not minimum variance) under a standard assumption of normality. More efficient estimates can be obtained using the full information maximum-likelihood (FIML) approach, which can be expressed as follows:

$$\begin{aligned} \ln L = \sum_{d_i=0} \ln \Phi(-\alpha'z_i) + \sum_{d_i=1} \left[-\ln \sigma_\varepsilon + \ln \phi \left(\frac{y_i - \beta'x_i}{\sigma_\varepsilon} \right) \right. \\ \left. + \ln \Phi \left(\frac{\alpha'z_i + \rho\sigma_\varepsilon^{-1}(y_i - \beta'x_i)}{\sqrt{1 - \rho^2}} \right) \right]. \end{aligned}$$

Maximizing this likelihood function produces simultaneous estimates of the parameters of both the discrete and continuous dimensions (α , β , ρ , and σ_ε). If $\rho = 0$, the log likelihood function reduces to the sum of a probit and a standard OLS regression, which can each be estimated separately. In comparison to the two-step Heckman procedure described above, the FIML estimator is computationally intensive to numerically identify optimal values. Reasonable starting values for FIML that are close to the true parameter values are therefore required. In this study, the final values of the Heckman two-step estimation procedure were used as the starting values for the FIML procedure.

4 Estimation results

Prior to undertaking estimation of the statistical model, the degree of correlation between each pair of the candidate independent variables was calculated. All such correlation coefficients were found to be smaller than 0.40; it was therefore determined that explicit correction for multicollinearity was not necessary.

A structured specification search was then undertaken, using the standard Akaike Information Criterion (AIC) to select between alternative candidate specifications. AIC is

² Note that the estimator of σ_ε can be computed using the procedure described in Step 1, which yields an estimate of the correlation coefficient is obtained: $\hat{\rho} = \hat{\beta}_\lambda / \hat{\sigma}_\varepsilon$.

Table 2 Estimation results

	Move/do not move dimension		Distance between LRT station and new residence	
	Coefficient	<i>t</i> statistic	Coefficient	<i>t</i> statistic
Constant	-1.44	8.72	-1.45	7.81
Household income				
Up to \$35 K	Reference class	N/A	Reference class	N/A
\$35–100 K	-0.185	2.25	-0.391	7.15
\$100 K+	-0.458	4.01	-1.231	5.12
Age				
Up to age 34	Reference class	N/A	Reference class	N/A
35–55	-0.135	1.48	0.213	1.64
55+	-0.506	3.15	0.439	7.32
Household Size				
One or two members	Reference class	N/A	Reference class	N/A
Three or more members	-0.185	1.67	0.432	6.40
Mode used to access LRT				
Drive	Reference class	N/A	Reference class	N/A
Transit	0.496	3.94	1.231	4.35
Walk	0.427	4.03	-0.644	5.33
Number of cars owned by household	-0.103	2.33	1.219	3.12
Number of children (under age 18) in household	-0.268	2.31	2.321	2.43
Length of time LRT has been used (months)	0.556	1.51	0.149	2.41
		σ_ε	0.81	2.34
		ρ	0.92	2.12
		Wald Chi square		56.73 ($p < 0.001$)

a metric of global goodness-of-fit which penalizes added parameters, and is widely used as to determine objectively whether the improved goodness-of-fit due to adding additional free parameters is warranted on the grounds of information theory [12].

The preferred model specification and resulting parameter estimates are presented in Table 2.³ Positive coefficients in the *move/do not move* dimension indicate that the relevant variable has a positive effect, *ceteris paribus*, on the likelihood of a survey respondent having relocated to the HBLR corridor within the prior 5 years and likewise positive coefficients in the *distance between LRT station and new residence* dimension are interpreted to mean that this distance increases, *ceteris paribus*, with the value of that variable.

As shown in Table 2, the majority of parameter estimates are statistically significant at the $p = 0.05$ level of confidence. The signs of parameter estimates are consistent with a priori expectations, as discussed in the remainder of

this section. Further, the structural parameters are both estimated to be statistically different than 1.0 ($t = 2.12$, 2.34, respectively), indicating the presence of statistically significant correlation between the error terms of the discrete and continuous models.

Household income was found to relate negatively, net of confounding effects, with the likelihood of having relocated in the prior 5 years; in other words having a lower household income was associated with a lower level of residential stability. Likewise, having a higher household income was associated with shorter distances between the new residences of residential-movers and the nearest LRT station. The fact that higher income groups chose to live closer to the stations of the LRT network is consistent with access to the LRT being a normal good, though this analysis cannot determine this definitively. Likewise, this result is also consistent with bid-rent theory, in which locations with higher accessibility are allocated to the land use that is willing/able to pay the most to occupy them [13].

Being under age 35 was associated with the highest propensity to have relocated within the prior 5 years, and also with the greatest propensity to relocate near to the access points to the LRT system. These findings are consistent with a more general pattern of younger adults

³ Variables that were tested and excluded from the preferred model specification, on the basis of AIC values, were as follows: dwelling type, race, employment status, occupation type, destination-type (of the surveyed journey), and whether or not the surveyed passenger receives a fare subsidy through their employer.

having greater residential mobility (cf. [14]), and also being more sensitive to the accessibility afforded by the LRT system. Similarly intuitive effects were found with respect to household size and the presence of children.

Automobile ownership was, by contrast, negatively associated with propensity to have relocated and, conditional on having moved, the proximity to the nearest LRT station. Having driven to access the LRT system for the surveyed journey was also negatively linked with having relocated (relative to having walked or taken another form of public transit). However, conditional on having relocated the largest distance between new home and nearest LRT station was associated with having used transit to access the LRT system—longer distance than both driving and walking.

Finally, Table 2 shows that the length of time that one has been using the LRT system is not significantly associated with the propensity to have relocated within the prior 5 years ($t = 1.51$), though is positively associated with the distance between [new] home and LRT station. Put another way, this suggests that people who started using the LRT service more recently are likely to live closer to the LRT system's stations than people who have been using it for a longer period of time (all else equal).

5 Conclusions

This paper contributes to the body of literature regarding the effects of public transport investments on residential property markets. Specifically, we investigated the effects of a new LRT system (the HBLR system in New Jersey, US) on: (1) the likelihood of LRT passengers having relocated in the prior 5 years and (2) the proximity of the 'new' residential location to the LRT system's stations. Given our a priori expectation of correlation in unobserved effects across these two dimensions of analysis, we employed a simultaneous discrete–continuous specification (a Type II Tobit). The empirical results suggest that the residual error terms were indeed correlated across these two dimensions; failing to take this correlation into account would have yielded biased and inefficient parameter estimates (i.e., effects).

The substantive results indicate that, among the sample of surveyed LRT passengers ($n = 1,023$), small household size, low income, being a young adult (under age 35) and low household car ownership are each independently associated with heightened propensity to have residentially relocated during the prior 5 years. Subsequently, conditional on having relocated, these same characteristics (with the exception of low income) are associated with relocating in close proximity to the LRT network's access points (stations).

Further research will be required to establish whether the empirical findings reported here are indicative of

generally applicable relationships, or are idiosyncratic to the HBLR system and/or the dynamics of the local real estate market in which it is located. For instance, the HBLR serves in part as a connecting service to heavy rail and ferry services to Manhattan's labor market and cultural destinations. It would be worthwhile, for instance, to establish whether or not similar functional relationships hold for LRT systems that comprise the core of their urban area's transit network (e.g., Portland Oregon's MAX system or Calgary's C-Train network), and for urban areas with characteristics different than the New York/New Jersey metropolitan region.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Cervero R, Landis J (1997) Twenty years of the bay area rapid transit system: land use and development impacts. *Transp Rev* A 31(4):309–333
2. Guiliano G (1989) New directions for understand transportation and land use. *Environ Plan A* 21:145–159
3. Robins ME, Wells JS (2008) Land development at selected Hudson–Bergen Light Rail Stations. Alan M. Voorhees Transportation Center (VTC) at Rutgers. April, 2008
4. Cervero R, Ferrell C (2002) Murphy S Transit-oriented development and joint development in the United States: a literature review. TCRP Report #52. Transportation Research Board, October 2002
5. Lund H (2006) Reasons for living in a transit-oriented development, and associated transit use. *J Am Plan Assoc* 72:357–366
6. Ketraungroch P (2008) The effect of mass rapid rail transit on residential location pattern in Bangkok. Master Thesis, Thammasat University, Bangkok, Thailand
7. Olaru D, Smith B, Taplin JHE (2011) Residential location and transit-oriented development in a new rail corridor. *Transp Res Part A Policy Pract* 45(3):219–237
8. Cao J, Schon J (2013) Transportation impact of transitways: a case study of Hiawatha light rail transit in Minneapolis. CTS Report 13–13, Report #7 in the series Transitway Impacts Research Program
9. Dai D, Weinzimmer D (2014) Riding first class: impacts of silicon valley shuttles on commute and residential location choice. Working Paper, UCB-ITS-WP-2014-01, February 2014
10. Liu R (2009) Hudson Bergen Light Rail customer impact and retention study. Final Report
11. Heckman JJ (1979) Sample selection bias as a specification error. *Econometrica* 47:153–161
12. Greene WH (2012) *Econometric analysis*, 7th edn. Prentice Hall, New York
13. Alonso W (1964) *Location and land value*. Harvard University Press, Cambridge
14. Schachter J (2001) Geographical mobility population characteristics March 1999–March 2000. US Census Bureau