Identifying Wider Barrier Effects of Ground-Level Transport Infrastructure through Spatial Causal Inference

Youngwoong Kim[1], Chaeyoung Shin[2], Sungmin Bae[2] Dongwoo Lee*[3]

[1] Ph.D. Student, Department of Smart Cities, University of Seoul, South Korea

[2] M.S. Student, Department of Smart Cities, University of Seoul, South Korea

[3] Associate Professor, Department of Smart Cities, University of Seoul, South Korea

SHORT SUMMARY

This study examined the wider barrier effects generated by ground-level urban expressways and railways. To address spatial confounding, spatial dependence, and spatial spillover in a cross-sectional causal-inference design, this study developed a spatial causal-inference framework that combines a multi-treatment design with Double Machine Learning and Gaussian-process-based nuisance learning. Using Seoul's ground-level transport networks and officially assessed land prices, the analysis found that the direct effects of transport infrastructure were mostly insignificant, whereas spillover effects were significant in most areas and reduced officially assessed land prices by an average of 11.12%. The results also showed that land-value declines were particularly pronounced in peripheral and relatively marginalized areas of Seoul, suggesting that transport infrastructure accelerates decline in already disadvantaged neighborhoods.

Keywords: Barrier effects, Gaussian process, Spatial causal inference, Transport equity

1. INTRODUCTION

Transport infrastructures such as highways, railways, and arterial roads can improve accessibility within and between regions and cities. At the same time, transport infrastructures can create physical barriers that disrupt the internal continuity of urban areas. Physical barriers may hinder access to everyday activities, increase travel costs, weaken local connectivity, and reduce opportunities for social interaction among neighbors. When such effects accumulate over time, physical barriers can contribute to community severance, distort nearby land use, and ultimately induce localized urban decline. Previous studies have referred to these unintended urban decline effects of transport infrastructure as barrier effects (Anciaes et al., 2016; van Eldijk et al., 2022)

Previous studies have classified barrier effects into direct, indirect, and wider barrier effects according to the affected domain and the scope of impact (van Eldijk et al., 2022). Direct and indirect barrier effects refer to changes at the level of individual behavior, such as increased travel costs and altered travel patterns caused by physical severance. Wider barrier effects refer to processes through which these mobility changes spread from individuals to communities and, ultimately, to society as a whole. Wider barrier effects associated with transport networks may isolate neighborhood commercial districts, weaken local accessibility, and reduce land values, thereby contributing to urban decline. For this reason, wider barrier effects have drawn increasing attention in transport and urban research.

Although theoretical discussion of wider barrier effects has expanded, empirical evidence based on causal identification remains limited. Causal identification is difficult for at least three

reasons. Fine-grained longitudinal data collected before and after infrastructure construction are rarely available in older urban settings. Transport infrastructure is not randomly located in urban space, so pre-existing locational conditions and latent spatial structure can confound the relation between barriers and urban outcomes. Wider barrier effects also tend to extend beyond directly exposed areas, which creates spatial spillovers and violates no-interference assumptions. Spatial causal-inference research has shown that spatial dependence, spatial heterogeneity, and spatial interference can bias conventional causal estimators when those features are ignored(Akbari et al., 2023; Bardaka et al., 2019). Transport studies have likewise shown that multiple transportation interventions and spillover effects require designs that move beyond a single binary treatment.

Against this background, the present study develops a spatial causal-inference framework tailored to urban barrier effects. Barrier exposure is defined as a multi-treatment structure that distinguishes direct exposure from nearby spillover exposure, rather than reducing transport infrastructure to a simple presence–absence condition. The framework then combines Double Machine Learning with Gaussian-process-based nuisance learning. Double Machine Learning supports flexible confounder adjustment through orthogonalization and reduces bias from regularized nuisance estimation(Chernozhukov et al., 2024; Wager & Athey, 2017). Gaussian-process-based nuisance learning captures latent spatial structure that is not fully represented by observed covariates(Rasmussen & Williams, 2006; Sigrist, 2022). This design therefore addresses three methodological challenges at the same time: endogenous infrastructure siting, unobserved spatial confounding, and spillover effects that violate no-interference assumptions.

The proposed framework contributes to the literature in both methodological and empirical terms. Methodologically, the proposed framework links multi-treatment causal design, interference-aware treatment definition, and spatially structured nuisance learning within a single estimation framework. Empirically, the proposed framework provides grid-level causal evidence on wider barrier effects, a topic that has so far remained largely at the level of theoretical discussion and descriptive interpretation. The resulting evidence helps clarify how transport-network severance contributes to localized urban decline in a dense metropolitan setting

2. RESEARCH DESIGN

Research scope

This study examines the barrier effects of ground-level urban expressways and railway networks across Seoul. To distinguish these effects from severance caused by natural environments, transport networks located adjacent to rivers and mountainous areas were excluded from the analysis. Barrier effects generated by transport infrastructure occur locally around the infrastructure itself, and analyses based on administrative districts may underestimate the barrier effects that arise within such areas. Accordingly, to quantify wider barrier effects at a finer spatial scale, this study conducted the analysis using 500 m grid cells. The spatial scope of the study is shown in **Figure 1**.

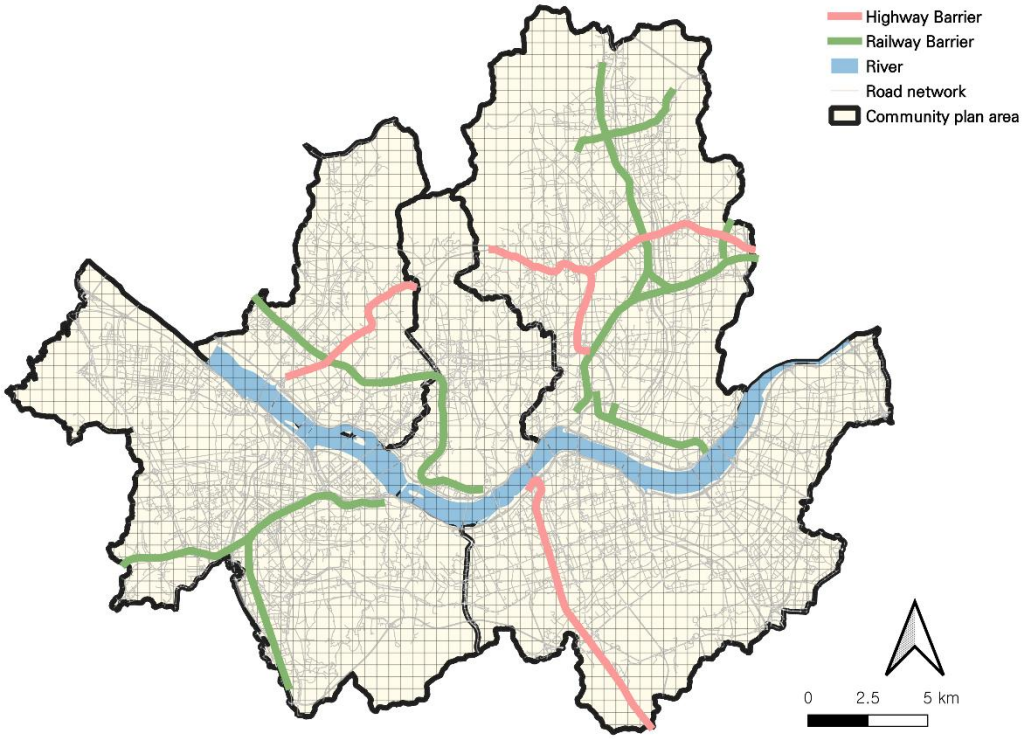


Figure 1. Research scope

Identifying the causal effects of transport-network barriers ideally requires longitudinal data collected before and after infrastructure construction. In Seoul, however, most ground-level transport networks were constructed during the twentieth century, which makes it extremely difficult in practice to obtain pre-construction grid-level data. For this reason, this study estimates barrier effects using cross-sectional data collected for the year 2023.

Variable selection strategy for causal inference identification

Officially assessed land price is a critical indicator of wider barrier effects because land price capitalizes the long-term consequences of accessibility, neighborhood interaction, development potential, and spatial disamenities into urban space. This analysis therefore uses officially assessed land price as a representative indicator of wider barrier effects and urban decline associated with transport-network severance. Since officially assessed land price showed substantial positive skewness and high kurtosis, a log transformation was applied to improve distributional normality.

In cross-sectional causal inference, control-variable selection should follow causal structure rather than statistical convenience, since adjustment for mediators or colliders can bias causal estimates through overcontrol or endogenous selection (Elwert & Winship, 2014; Hernán et al., 2002)Wider barrier effects, in particular, arise through the influence of direct and indirect barrier effects, such as reduced accessibility(Anciaes et al., 2016; van Eldijk et al., 2022). Controlling for direct and indirect effects in the estimation of wider barrier effects may therefore induce overcontrol bias.

Accordingly, this study controlled for locational and environmental characteristics that may affect land values, including distance to the urban center, land use, and average building height, drawing on the logic of the hedonic price model. By contrast, accessibility-related variables that may themselves have been affected by direct and indirect barrier effects were treated as mediators and were therefore not controlled for. Network density and crosswalk count were nevertheless included, because previous barrier-effect studies have identified street-network structure and crossing facilities as key determinants of the magnitude of severance. In addition, to distinguish the accessibility benefits of rail station proximity from the severance effects of ground-level railways, this study controlled for a dummy variable indicating the presence of a subway station.

Table 1: Descriptive Statistics of Variable

	Variable	**Mean**	**S.D**	**Min**	**Max**
Covariates	Population	5,067	3,376	16	17,925
	Company count	577	669	0	8,929
	Network Density	10	8.3	0	75.80
	Old building	227	218	0	1,177
	BD height	2,592	5,688	0	127,363
	Crosswalk Count	6.34	5.86	0	49
	Commercial area ratio	0.05	0.14	0	1.00
Confounding variable	Distance to CBD	10	3.25	5.25	19.09
	Average Slope	5.25	5.44	0	31.01
	Rail node dummy	0.03	0.18	0	1
Dependent variable	Log transformed Land price	14.94	0.970	10.41	17.63

Wider barrier effects can extend beyond areas directly exposed to the barrier and may generate spillover effects in nearby areas, thereby violating the Stable Unit Treatment Value Assumption (SUTVA). For this reason, analyzing wider barrier effects requires simultaneous consideration of both the effect on directly exposed areas and the spillover effect that propagates into surrounding areas. To account for the spillover effect of barrier impacts, this study defines treatment status using a multi-treatment design, as shown in Equation (1).

$$\mathbf{T}_i = \begin{bmatrix} D_i \\ S_i \end{bmatrix} \tag{1}$$

Here, D_i indicates whether grid $i$is directly treated, and S_i denotes spillover exposure, measured as the proportion of directly treated grids among neighboring grids located within a 500 m radius. The resulting spatial distribution of the direct treatment group, the spillover treatment group, and the non-treated group is presented in **Figure 2.**

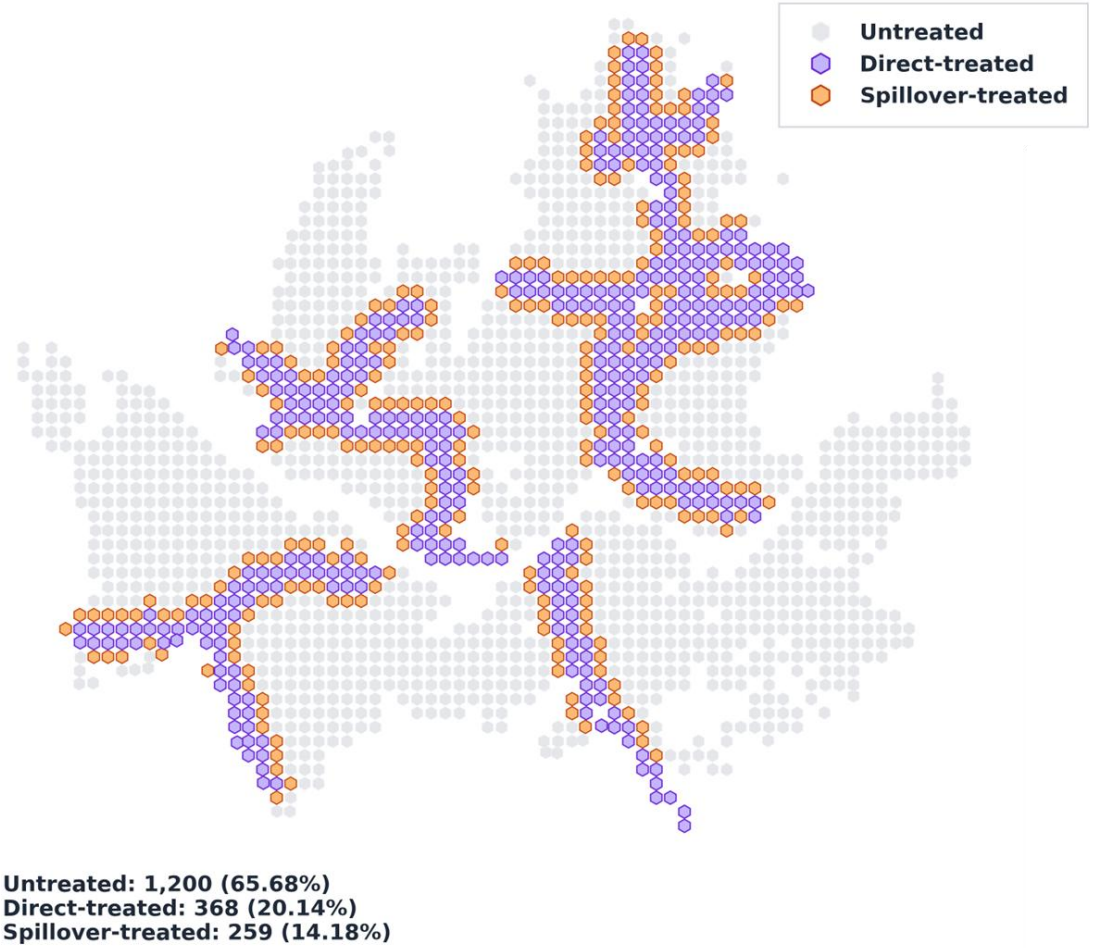


Figure 2. Spatial distribution of direct treatment, spillover treatment, and untreated grids

3. METHODOLOGY

Gaussian process augmented multi-treatment DML framework

The location of transport infrastructure and the surrounding spatial structure are not formed randomly. Factors such as accessibility to central areas and local environmental conditions can simultaneously affect not only the outcome variable but also both direct treatment and spillover exposure. Some of these factors, moreover, may be only partially observed or may exhibit spatially structured patterns, meaning that nuisance estimation based solely on the conventional IID assumption may not be sufficient. To mitigate this problem, this study combines a multi-treatment CausalForestDML (Chernozhukov et al., 2024; Wager & Athey, 2017) framework, which simultaneously incorporates direct treatment and spillover treatment, with a Gaussian Process (GP) design that absorbs latent spatial confounding at the nuisance-estimation stage.

This study extends the partially linear regression structure of DML proposed by Chernozhukov et al. (2024) to a design that simultaneously considers direct treatment and spillover treatment. The baseline specification is given in Equation (2):

$$Y_i = \theta(X_i)^\top \mathbf{T}_i + \mu_Y(X_i, W_i) + g_Y(s_i) + \epsilon_i \tag{2}$$

Here, $\mu_Y(X_i, W_i)$ represents the component explained by the observed covariates, whereas $g_Y(s_i)$ denotes an unobserved latent factor that varies across grid cells according to the spatial coordinate s_i. The spatial term $g_Y(s_i)$ is modeled as in Equation (3):

$$g_Y(\cdot) \sim GP(0, K_y(\cdot,\cdot)) \tag{3}$$

In this expression, $K_Y(\cdot,\cdot)$ is a covariance function that captures spatial similarity, allowing unobserved spatial latent factors to be more similar for nearby observations. Depending on the form of the covariance function, nonlinear similarity structures between neighboring observations can also be accommodated (Rasmussen & Williams, 2006). In this study, exponential, Matérn, and Gaussian kernels were compared, and the exponential kernel was selected because it yielded the lowest prediction error after parameter tuning.

Outcome and treatment residualization by nuisance functions

In this study, nuisance functions for both the outcome and the treatment include observed fixed effects and GP-based spatial effects.

$$m_Y(X_i, W_i, s_i) = E[Y_i | X_i, W_i, s_i] \tag{4}$$

$$E[Y_i | X_i, W_i, s_i] = \mu_Y(X_i, W_i) + g_Y(s_i) \tag{5}$$

$$\widetilde{Y}_i = Y_i - \widehat{m}_y(X_i, W_i, s_i) \tag{6}$$

The residual $\tilde{Y}_i$ obtained from Equation (6) represents the variation in the outcome that is not explained by the fixed-effect component $\mu_Y(X_i, W_i)$ or by the unobserved spatial effect $g_Y(s_i)$.

The same orthogonalization procedure is then applied to the treatment vector. As the treatment vector in this study includes both direct treatment and spillover treatment, the treatment nuisance function is defined in vector form:

$$\boldsymbol{m}_T(X_i, W_i, s_i) = E[\mathbf{T}_i | X_i, W_i, s_i] \tag{7}$$

$$E[\mathbf{T}_i | X_i, W_i, s_i] = \mu_T(X_i, W_i) + g_T(s_i) \tag{8}$$

$$\mu_T(X_i, W_i) = \begin{bmatrix} \mu_D(X_i, W_i) \\ \mu_S(X_i, W_i) \end{bmatrix}, g_T(s_i) = \begin{bmatrix} g_D(s_i) \\ g_S(s_i) \end{bmatrix} \tag{9}$$

$$\widetilde{\mathbf{T}}_i = \mathbf{T}_i - \widehat{\boldsymbol{m}}_T(X_i, W_i, s_i) \tag{10}$$

Identifying heterogenous treatment effects by causal forest

Finally, this study estimates the function $\theta(X_i)^\top \mathbf{T}_i$, which explains the relationship between the two residuals, by applying a causal forest in order to analyze heterogeneous treatment effects, as shown in Equation (11).

$$\widetilde{Y}_i = \hat{\theta}(X_i)^\top \widetilde{\mathbf{T}}_i + u_i \tag{11}$$

Here, $\theta(X_i) = (\theta_d(X_i), \theta_s(X_i))^\top$is a vector of location-specific heterogeneous treatment effects that includes both the direct effect and the spillover effect. $\theta_d(X_i)$denotes the function describing the direct treatment effect when a given grid is directly exposed to the barrier, whereas $\theta_s(X_i)$represents the localized spillover effect through which barrier exposure in surrounding areas propagates to that grid. This framework therefore makes it possible to examine, on a location-specific basis, whether the barrier effect is confined to adjacent grids or extends into surrounding neighborhoods in the form of a wider barrier effect.

4. RESULTS AND DISCUSSION

Model performance was evaluated using the EconML score provided by CausalforestDML (Wager & Athey, 2017). The EconML score measures, in terms of mean squared error (MSE), how well the heterogeneous treatment effect explains the variation attributable to the treatment effect. The EconML score is calculated as shown in Equation (12).

$$\frac{1}{n}\sum_i^n (\widetilde{Y}_i - \hat{\theta}(X_i) \cdot \widetilde{\mathbf{T}}_i)^2 \tag{12}$$

To assess the extent to which the GP component controls for spatial autocorrelation and latent effects, this study compares the EconML score and Moran's I statistics of residualfrom GPBoostCF with that from RandomForestCF. The results are reported in **Table 2**.

Table 2: Goodness of fit statistics of model

	Causalforest with Gaussian Process (GP-CF)	**Causalforest with Random Forest (RF-CF)**
EconML Score	0.01	0.15
Moran's I Statistics of Residual	0.28***	0.57

*P<0.1, **P<0.05, ***P<0.01

The results show that the EconML score of the Causal Forest with Gaussian Process (GP-CF) was 0.010, whereas that of the Causal Forest with Random Forest (RF-CF) was 0.155, indicating that causal effects were estimated more accurately when spatial confounding was controlled through a probabilistic random-effect structure. Residual Moran's I statistics were statistically significant in both models, indicating that spatial autocorrelation still remained in the residuals. Even so, the Moran's I statistic of the GP-CF was 0.28, which was lower than that of the RF-CF (0.57), suggesting that the GP-CF partially alleviated spatial autocorrelation.

The results of the grid-level heterogeneous treatment effect analysis are presented in **Figure 3**. The mean direct treatment effect was estimated at -0.75%, and it was statistically significant in only 4% of the treated areas. By contrast, the mean spillover effect was estimated at -11.12%, and 31.74% of the treated areas were found to be statistically significant. The fact that the direct treatment effect was generally insignificant, whereas the spillover effect appeared stronger and more spatially extensive, may reflect the possibility that in areas directly traversed by ground-level transport networks, the disamenities caused by severance are capitalized together with benefits associated with nodal accessibility, visibility, and location along major corridors, thereby offsetting the net effect (Bae et al., 2003). Adjacent areas, by contrast, may receive fewer of the direct accessibility benefits provided by the transport corridor while simultaneously accumulating the direct and indirect disadvantages of reduced accessibility and neighborhood severance. This accumulation may then extend into a wider barrier effect (Anciaes et al., 2016; Gbban et al., 2023; van Eldijk et al., 2022), resulting in a stronger negative effect of the barrier on land values.

A particularly pronounced decline in land values was identified in southwestern Seoul. Southern Seoul contains a mixture of low-rise residential neighborhoods, semi-industrial districts, and aging housing environments, and it has relatively weak neighborhood-level connectivity within the city. For this reason, the spillover effects of infrastructural barriers may have accumulated more strongly there, leading to a more pronounced decline in land values.

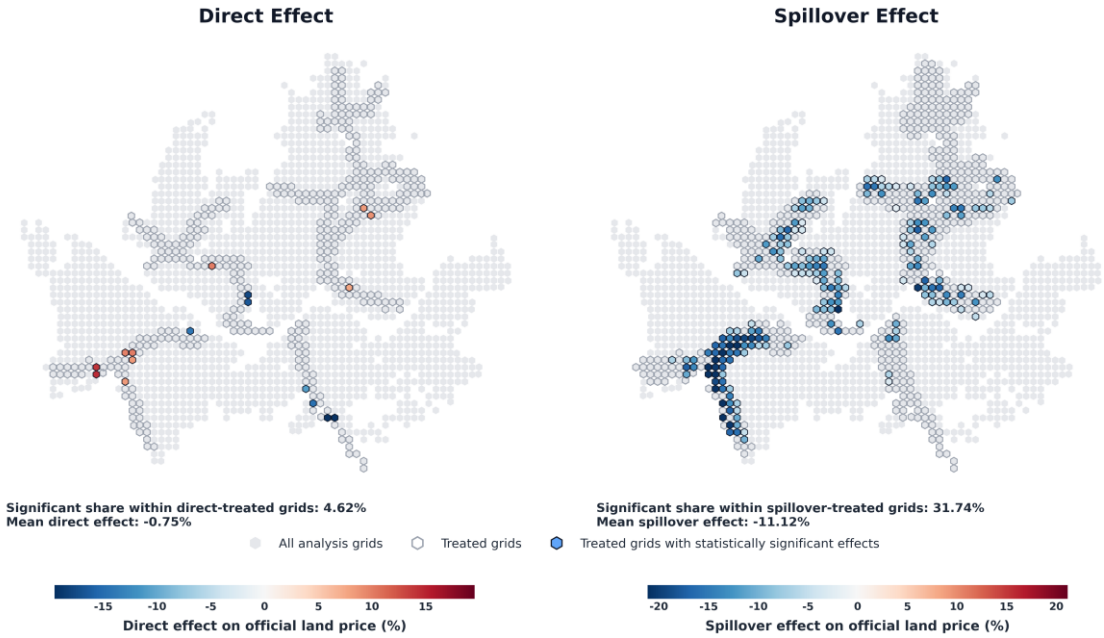


Figure 3. Spatial distribution of estimated direct and spillover effects on land price (Left panel: direct effects, Right panel: spillover effects)

5. CONCLUSIONS

This study provides causal evidence that ground-level transport infrastructure generates wider barrier effects on land price. The proposed framework combines a multi-treatment DML design with Gaussian-process-based nuisance learning to address spatial spillover, endogenous infrastructure siting, and unobserved spatial confounding in a cross-sectional setting. The results show that direct effects are generally weak, whereas spillover effects are significant in many treated areas and reduce officially assessed land prices by an average of 11.12%. The strongest negative effects appear in peripheral and relatively marginalized parts of Seoul, especially in the southwest, which suggests that transport-network severance can accelerate localized urban decline and deepen spatial inequality.

This study nevertheless has an important limitation in that it was unable to control for noise, air quality, and related externalities, and therefore could not fully separate the effects of physical severance caused by transport infrastructure from the effects of environmental externalities. Once data capable of quantifying such externalities, including noise and air-quality conditions, become available, future research will be able to analyze more precisely the decline effects attributable to transport-infrastructure severance.

ACKNOWLEDGEMENTS

This work was supported by the National Research Foundation of Korea (NRF-RS-2026-25498404)

REFERENCE

Akbari, K., Winter, S., & Tomko, M. (2023). Spatial Causality: A Systematic Review on Spatial Causal Inference. *Geographical Analysis*, *55*(1), 56–89. https://doi.org/10.1111/gean.12312

Anciaes, P. R., Jones, P., & Mindell, J. S. (2016). Community Severance: Where Is It Found and at What Cost? *Transport Reviews*, *36*(3), 293–317. https://doi.org/10.1080/01441647.2015.1077286

Bae, C.-H. C., Jun, M.-J., & Park, H. (2003). The impact of Seoul's subway Line 5 on residential property values. *Transport Policy*, *10*(2), 85–94. https://doi.org/10.1016/S0967-070X(02)00048-3

Bardaka, E., Delgado, M. S., & Florax, R. J. G. M. (2019). A spatial multiple treatment/multiple outcome difference-in-differences model with an application to urban rail infrastructure and gentrification. *Transportation Research Part A: Policy and Practice*, *121*, 325–345. https://doi.org/10.1016/j.tra.2019.01.028

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2024). *Double/Debiased Machine Learning for Treatment and Causal Parameters* (arXiv:1608.00060). arXiv. https://doi.org/10.48550/arXiv.1608.00060

Elwert, F., & Winship, C. (2014). Endogenous selection bias: The problem of conditioning on a collider variable. *Annual Review of Sociology*, *40*, 31–53. https://doi.org/10.1146/annurev-soc-071913-043455

Gbban, A. M., Kamruzzaman, Md., Delbosc, A., & Coxon, S. (2023). The wider barrier effects of public transport infrastructure: The case of level crossings in Melbourne. *Journal of Transport Geography*, *108*, 103554. https://doi.org/10.1016/j.jtrangeo.2023.103554

Hernán, M. A., Hernández-Díaz, S., Werler, M. M., & Mitchell, A. A. (2002). Causal Knowledge as a Prerequisite for Confounding Evaluation: An Application to Birth Defects Epidemiology. *American Journal of Epidemiology*, *155*(2), 176–184. https://doi.org/10.1093/aje/155.2.176

Rasmussen, C. E., & Williams, C. K. I. (2006). *Gaussian processes for machine learning*. MIT Press.

Sigrist, F. (2022). *Gaussian process boosting | The Journal of Machine Learning Research*. https://dl.acm.org/doi/abs/10.5555/3586589.3586821

van Eldijk, J., Gil, J., & Marcus, L. (2022). Disentangling barrier effects of transport infrastructure: Synthesising research for the practice of impact assessment. *European Transport Research Review*, *14*(1), 1. https://doi.org/10.1186/s12544-021-00517-y

Wager, S., & Athey, S. (2017). *Estimation and Inference of Heterogeneous Treatment Effects using Random Forests* (arXiv:1510.04342). arXiv. https://doi.org/10.48550/arXiv.1510.04342