The impact of automotive industry investments on local property prices – The case of Mercedes-Benz in Kecskemét (Hungary), 2010–2017

Nikolett Vágó
(corresponding author)
Central Bank of Hungary,
Financial System Analysis Directorate,
Budapest, Hungary
Email: vagon@mnb.hu

Sándor Winkler
Central Bank of Hungary,
Financial System Analysis Directorate,
Budapest, Hungary

Gábor Hajnal
Central Bank of Hungary,
Financial System Analysis Directorate,
Budapest, Hungary

The automotive industry, being one of the largest industries worldwide, is a dominant and increasingly important sector across Europe. Due to the large foreign direct investment inflow from multinational automotive companies, it has special importance in the Central and Eastern European region. Owing to their sheer size, large-scale greenfield investment projects by multinational carmakers considerably affect macroeconomic output as well as regional social and economic development, including the housing market. This study explores whether a large-scale investment project can cause regional housing prices to deviate from national trends. Using a multivariate synthetic control method estimator, the authors identify the extent to which local house prices were influenced by the Mercedes-Benz car factory investment in Kecskemét, Hungary. The study compare the evolution of house prices in Kecskemét to a synthetic control of other county seats. The results suggest that following the completion of the Mercedes factory, local house prices increased by 39 percentage points more between 2010 and 2017 compared to the counterfactual case. The estimated effect is significant and large in magnitude compared to the pseudo-effects based on spatial placebo tests. Several robustness analyses confirm this conclusion.

Keywords:
treatment effects,
synthetic control method,
regional heterogeneity,
automotive industry investment,
housing prices

Introduction

House prices have special relevance from several perspectives. Residential real estate is the main asset of households and serves as collateral for bank mortgage loans. Therefore, a change in housing prices not only affects the wealth of households, but also their creditworthiness and, indirectly, bank profitability by influencing losses if defaults happen. Housing-related costs are a significant part of household spending, and influence consumption and savings decisions. Moreover, residential real estate can serve as an investment vehicle; therefore, demand for it is also determined by its relative yield versus other investments. Finally, the housing market has significant economic policy relevance, as it is strongly connected to other parts of the economy. House prices play an important role in the economic cycle as well as in the feedback effects between the real economy and financial system.

This study is related to the literature on identifying the economic factors that explain price developments in the housing market. House prices are largely driven by demand-side fundamentals. The long-run equilibrium relationship between house prices and household income (Abraham–Hendershott 1996, Capozza et al. 2002), as well as employment growth (Baffoe-Bonnie 1998, Apergis–Rezitis 2003), is well established. Demographic factors[1], such as the change in and structure of the population or age composition also play an important role, especially in the long run. Further, ample empirical evidence exists on the relationship between credit availability and the evolution of property prices; however, conclusions about the direction of causality vary[2]. Interest rates are also associated with house prices and tend to move in opposite directions (Himmelberg et al. 2005, Taylor 2009, Agnello–Schuknecht 2011). Housing supply also plays a significant role in driving housing prices (Duca et al. 2021). An increase in construction costs has a long-term positive impact on house prices (Adams–Füss 2010, Xu–Tang 2014, Belke–Keil 2018), as well as supply constraints, such as land-use regulations and the physical lack of developable land (Glaeser et al. 2005, Quigley–Raphael 2005, Saks 2008, Glaeser et al. 2008, Hilber–Vermeulen 2016). Finally, the price effect of demand is greater with inelastic housing supply (Saiz 2010, Cavalleri et al. 2019).

[1] A larger population (Belke–Keil 2018, Takáts 2012) and decreasing average household size (Day 2018) can cause higher housing prices due to higher demand. Population structure and the age composition also matter. The relationship between old-age dependency ratio and house prices has been keenly investigated, albeit with conflicting results regarding the sign of the effect (Takáts 2012, Hiller–Lerbs 2016, Lee–Jung 2020, Wang–Kinugasa 2022). Rising life expectancy, international immigration, and urbanisation have a long-lasting positive effect on house prices, while declining fertility puts a downward pressure on them (Gong–Yao 2022).

[2] Fitzpatrick–McQuinn (2004) and Gimeno–Martínez-Carrascal (2006) find that the two variables are interdependent in the long run, Brissimis–Vlassopoulos (2009) report an absence of long run causation running from housing loans to housing prices, while Gerlach–Peng (2005) conclude that the direction of causality runs from property prices to credit.

This study explores the effect of the Mercedes-Benz car factory investment project in Kecskemét[3] on the value of local properties, investigating whether a large-scale investment project can put a region's property price developments on a path different from that of the national average. Owing to the steep and geographically heterogeneous rise in house prices seen in Hungary since 2014[4], attention has increasingly turned towards identifying the factors that determine house price appreciation. Besides general economic factors, regional elements also play a substantial role. For example, car factory investments may influence house prices in the affected settlements because of, among other things, the surplus demand by the workforce flowing into the cities because of the investment, the supplier network linked to it, and infrastructure developments. Property owners' wealth increases with the persistent rise in house prices. Simultaneously, the affordability of housing may deteriorate for the people moving in and the local population in the case of a lag in income development, which can negatively affect workforce mobility.

We apply a data-driven econometric estimator, the synthetic control method (SCM), proposed by Abadie–Gardeazabal (2003) and further developed by Becker–Klößner (2018). SCM is particularly useful for estimating causal effects with the panel data structure we have; that is, a single unit is treated and relatively few similar control units are available. Moreover, it offers a significant advantage over standard modelling practices in that the counterfactual of the treated unit is generated in a transparent, data-driven process approximated by a synthetic control, which is a weighted average of potential control units. These weights are determined in such a way that the difference between the synthetic control and treated units is minimised with respect to both the outcome variable and some covariates prior to the event. If the outcome variables of the treated units and the synthetic control vary substantially following an event, the estimated effect can be attributed to the event in question. The SCM also allows one to run placebo tests to evaluate the statistical significance of the estimated effect.

According to the estimation results, the treatment – that is, the completion of the Mercedes factory project – significantly modified the long-term development of local house prices. Goodness-of-fit indicators show that the estimated synthetic control time series fits well onto the house price index of the treated city in the pre-intervention period; therefore, it is an appropriate approximation of the unobservable counterfactual. Thus, the key identifying assumption of the SCM is fulfilled, and the model is suitable for inference. The synthetic control estimation suggests that the treatment substantially affected local house prices. In the period between 2010 Q4

[3] Kecskemét is the eighth largest city in Hungary, has a population of around 110,000, and is located a 60-minute drive south of Budapest.

[4] Based on the MNB house price index, house prices increased by 216% in Budapest, 163% in the country overall, and 156% in cities outside the capital in nominal terms between 2014 Q1 and 2021 Q2. For a detailed description of housing market developments (see second section).

and 2017 Q4, house prices increased 39 percentage points more in Kecskemét than they would have without the car factory investment. We test the significance of the effect using two indicators based on spatial placebo tests in which the treatment is iteratively applied to all other county seats. The estimated effect for Kecskemét is large in magnitude compared to the pseudo effects. Further, the treatment can be considered significant based on the ratio of post- to pre-treatment root-mean square percentage error (RMSPE) measures, as well as the relative gaps calculated for each period.

We contribute to the literature which quantifies the effects of investment projects on property prices. A particularly studied topic is the relationship between housing prices and changes in spatial accessibility. Several studies have focused on the house price effect of railway investments (Mohammad et al. 2013)[5]. Others have also conducted impact assessments of investments in road-based transportation infrastructure (Mikelbank 2004) and public transport (Dai et al. 2016, Béres et al. 2019, Tan et al. 2019). Most of these studies find a significant positive relationship between improved accessibility and house prices. Some studies also focus on the causal relationship between brownfield redevelopment and property price changes (De Sousa 2005, De Sousa et al. 2009, Haninger et al. 2017, Linn 2013, Woo–Lee 2016), and find evidence of positive spillover effects. A few authors have also analysed the extent to which foreign direct investments affect house prices. Agnew–Lyons (2018) examine the impact of changes in employment in internationally trading (foreign direct investment (FDI)) firms on housing prices in Ireland, while Kim–Lee (2022) analyse the impacts of FDI on house prices in China. Other studies have focused on the price effects of the negative externalities related to industrial facilities (Currie et al. 2015, Grislain-Letrémy–Katossky 2014, Boxall et al. 2005, de Vor–de Groot 2011).

The main contribution of this study is that, to the best of our knowledge, it is the first empirical study to provide evidence that a car factory investment project can substantially affect regional house prices[6]. Moreover, we identify the housing price effect of a large-scale investment project using a counterfactual method, namely, the multivariate synthetic control method estimator (MSCMT). Our results are particularly relevant because of the scale and importance of the automotive industry as well as its ongoing transformation due to the green transition. The automotive industry is one of the largest industries worldwide, employing approximately 4 million people [1]. It is also a dominant and increasingly important sector across Europe: in 2011, it accounted for 2.7% of the gross value added in the EU, which rose to 3.1% by 2018 [2]. Due to the large FDI inflow from multinational automotive companies,

[5] Numerous studies estimating the impact of rail investments on property prices are summarised in this meta-analysis.

[6] The effect of FDI inflow on the local economy of a Hungarian county seat was investigated in Antalóczy et al. (2022).

the significance of the industry is even higher in the Central and Eastern European (CEE) region[7]; foreign automotive lead firms invested more than EUR 35 billion between 1990 and 2015 (Pavlínek et al. 2017). Moreover, the transition to a sustainable economy requires large physical investments in Europe's automotive industry. So far, seven new battery factory investments with a capacity of 100 GWh have been planned for the 2020s. Overall, the construction of approximately 24 large-scale battery factories is necessary across Europe to cover the rising demand for electric vehicles (according to McKinsey's calculations [5]). The automotive industry and large-scale investment projects in the sector not only have macroeconomic significance but also substantially affect local social and economic developments, including property prices. Owing to its role in the economy[8], this is also the case in Hungary, which is the location of the treatment under review.

The remainder of this article is organised as follows. The second section describes the economic and property market developments in Hungarian county seats using the database employed for modelling. The third section details the methodology, while the fourth section discusses the model specifications and estimation results. The fifth section presents the robustness analyses. Finally, the sixth section summarises the main findings.

Data and a descriptive analysis of housing market developments in Kecskemét

We use quarterly and annual settlement-level data from 2000 to 2017, with 72 observation periods at a quarterly frequency. The database was narrowed down to 18 county seats in Hungary so that the synthetic control time series included settlements which are of similar importance to Kecskemét. The Hungarian capital, Budapest, was not included in the database, as it differs considerably from county seats in terms of size and economic significance.

The outcome variable was the house price index calculated for county seats, produced using the Central Bank of Hungary (MNB) house price index methodology detailed in Banai et al. (2017, 2018). While estimating the house price indices, county seats were supplemented with nearby settlements, which can be considered a part of the local housing market owing to their proximity to the county seat. Their inclusion

[7] In 2018, it represented 6.4% of gross value added and 4.1% of employment on average in the Visegrád countries and Romania, which is double the EU average.

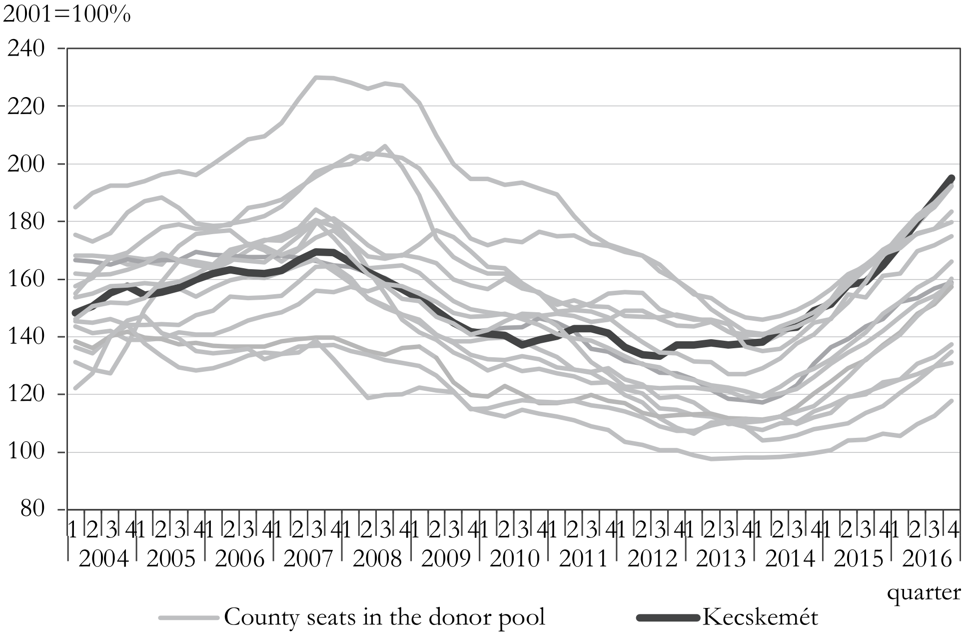
[8] In Hungary, 104,000 people, or 3.6% of the total workforce, were employed in the automotive industry in 2018. Further, the sector's contribution to gross value added was 6.6%. The automotive industry's special importance in the national economy is also reflected by the fact that among the currently ongoing construction projects, the battery factory in Iváncsa (EUR 1.85 billion) and the BMW factory in Debrecen (EUR 0.87 billion) are among the largest construction projects in Hungary in terms of the capital invested. Meanwhile, investments for a CATL battery factory in Debrecen are also being prepared.

also expanded the estimation sample of the indices, which improved the reliability of the estimates.

Figure 1 shows the house price indices for the county seats included in the estimate. The house price appreciation in Kecskemét, the focus of our analysis, was around the middle of the range of appreciation of other county seats until the 2008 economic crisis. Between 2008 and late 2013, house prices decreased steadily across Hungary. However, in Kecskemét, house prices stopped falling significantly after 2010, when the Mercedes factory was already underway; in fact, a stagnant trend, with some minor fluctuations, could be observed until early 2014. Overall, among the county seats reviewed, Kecskemét experienced the largest house price appreciation between 2001 and 2016. At the end of 2016, Kecskemét had the sixth highest average square metre prices, whereas it was ranked 11th in the early 2000s.

Figure 1

Estimated house price indices for the county seats



Source: [7] levy database and authors' calculations.

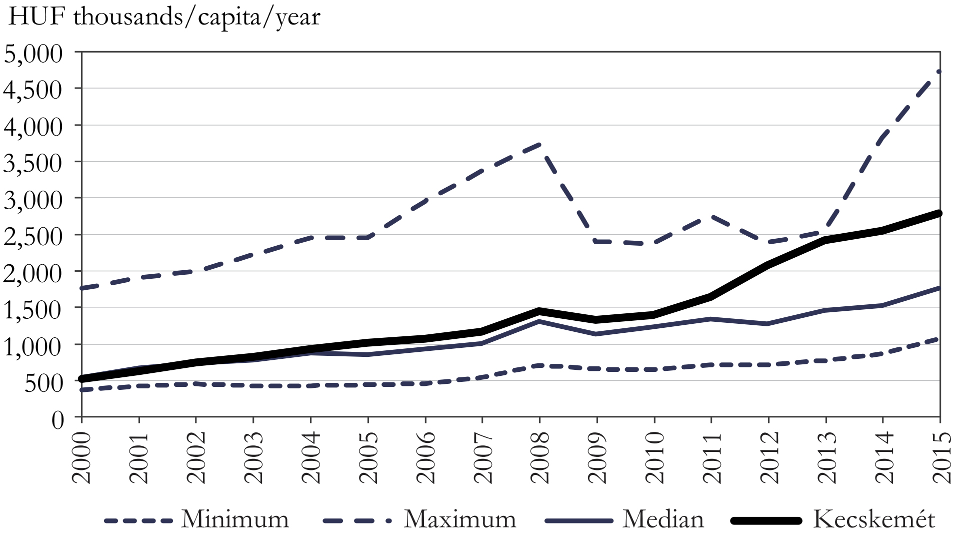
Note: The samples used to estimate the price indices include transactions in settlements near the respective county seats.

The covariates included in the estimation were selected to capture the long-run housing market conditions as well as the social and economic factors affecting such conditions. Variables describing the size and location of the cities are important; thus, the stock of dwellings, population, and shortest distance to Budapest were included in the estimates. Housing market conditions in the county seats were approximated by the ratio of the number of transactions to the stock of dwellings (i.e. the turnover rate), the average square metre prices, the ratio of new homes to the stock of dwellings, the ratio of empty homes to the stock of dwellings, and an indicator capturing the affordability of home purchases. The economic development of the settlements was captured using three indicators: per-capita income after tax, companies' aggregate equity, and corporate export sales revenue. Demographic trends in the settlements were represented by net migration per one thousand people. Finally, variables reflecting the civil and social dimensions of county seats were included: the ratio of employed people, ratio of community housing to the stock of dwellings, ratio of homes with modern conveniences to the stock of dwellings, and number of hospital beds per one thousand people. The definitions and time horizons of the covariates are listed in Table A1 of the Appendix, while Table A2 shows the summary statistics for the pre- and post-treatment periods.

The impact of the Mercedes factory in Kecskemét on the local housing market supply and demand is discussed below, along with the extent to which social and economic developments changed relative to other county seats after the investment project.

Figure 2

Per capita value added in Hungarian county seats



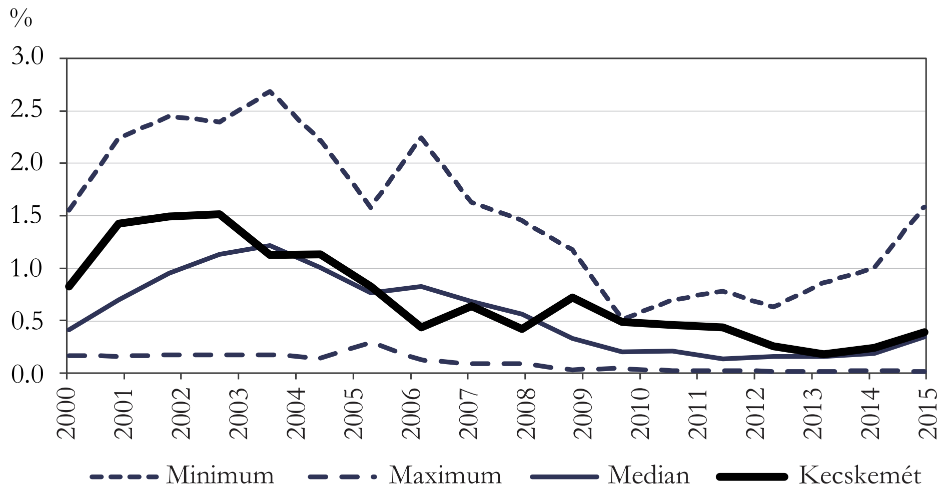
Source: [6], [4].

First, income conditions, as a fundamental feature of housing market demand, are examined based on the evolution of per capita value added[9] The value added produced by companies serves as a basis for wage setting. This indicator measures not only households' disposable income but also incorporates the income of smaller firms not living from earned income. While Kecskemét had the fifth lowest per capita gross value added among Hungarian county seats in 1998, after the Mercedes investment in 2011, income started rising sharply in the city, exceeding the median value of other county seats (Figure 2). By 2015, Kecskemét had the third-highest per capita gross value added. Moreover, after the opening of the Mercedes factory (2011–2012), Kecskemét differed from other county seats in its demographic developments: 2011 saw the highest net domestic migration per one thousand people (the difference between those moving in and out) among county seats. Besides the growing income, the additional housing market demand generated by the factory's labour demand may also have played a role in the rising property prices in Kecskemét.

Regarding housing market conditions, another important question is how the housing market supply responded to a large-scale project like the construction of the Mercedes factory. The Mercedes investment in Kecskemét was implemented when the financial crisis spilled over to Hungary and spread throughout the economy. During this period, Hungarian housing market activity continued to decline. This was reflected not only in the falling number of transactions and house prices, but also in the number of homes built. While 32,000 new homes were completed in Hungary in 2009, the number of completions continued to shrink steadily until 2013, when the market reached its trough, at 7,300 completions across the country. However, the renewal of the stock of dwellings in different county seats considerably varied. The annual renewal rate of the stock of dwellings in Kecskemét was 1.5% in the early 2000s before plummeting to below 0.5% by 2007, which marked a lower home construction ratio than that in most county seats. However, two years after the announcement of the Mercedes factory in 2008, the renewal rate of the stock of dwellings in Kecskemét increased, while all other county seats experienced declines in new home construction owing to the property market slump. Until the opening of the factory in 2012 and two following years, the renewal rate in Kecskemét was substantially higher than the median renewal of county seats, suggesting that investors responded to housing demand in the city (Figure 3).

[9] This variable was excluded from the estimations due to the shorter length of the time series. However, it represents the structural changes in Kecskemét after the Mercedes investment well enough.

Figure 3
Annual renewal rate of the stock of dwellings (number of homes constructed relative to the stock of dwellings)



Source: [6], [4].

Methodology

We use the SCM[10] to assess whether the Mercedes factory investment in Kecskemét significantly affected house prices in the city and its neighbouring settlements, and identify the extent of the impact on home values. In the Rubin causal model's conceptual framework (Imbens–Rubin 2015), the Mercedes factory investment can be considered as the treatment. Meanwhile, the SCM is intended to approximate the unobservable counterfactual time series[11], or the extent to which house prices would have deviated in the post-treatment period without the treatment event.

SCM was first used by Abadie–Gardeazabal (2003) to assess the impact of attacks by the Basque Homeland and Liberty (ETA) terrorist organisation on the local economy. Since the publication of Abadie et al. (2010), this method has been applied in several disciplines to estimate the effects of various events, such as natural disasters (Barone–Mocetti 2014, Coffman–Noy 2012), economic liberalisation (Billmeier–Nannicini 2013), historically important economic or political decisions (Abadie et al. 2015, Campos et al. 2019, Opatrny 2021), mass migration (Peri–Yasenov 2019), and economic policies (Eren–Ozbelik 2017, Falkenhall et al. 2020). Several studies have

[10] The reasons for using the SCM for answering our research question is detailed in the Appendix (1. Advantages of the SCM).

[11] We elaborate on the identifying assumptions related to SCM in the Appendix (2. Identifying assumptions).

used SCM to estimate the impact of an event on housing prices (Gautier et al. 2009, Bronzini et al. 2020, Gabriel et al. 2021, Rouwendal–Petrat 2022).

Furthermore, the standard SCM method has been expanded and enhanced in recent years based on various considerations[12]. In response to the estimation bias that arises when only the average value of economic predictors is used to generate the weights for the synthetic control, Becker–Klößner (2018) developed the MSCMT. This method can account for all pre-treatment covariate values while creating the control group; thus, the information content in the change in economic predictors over time can be utilised when establishing the counterfactual state. Another advantage of the MSCMT algorithm is that it yields more reliable and accurate results, and allows one to simultaneously examine multiple outcome variables. This method has been used in several analyses, including Grabovac et al. (2018), Martinelli–Vega (2019), and Monastiriotis–Zilic (2020).

Programme evaluation with the synthetic control method

The impact of the factory investment in Kecskemét on local house prices is derived from the difference between the actual and counterfactual time series, $\beta_t = Y_{1t} - \tilde{Y}_{1t}$, where Y_{1t} is the actual value of the Kecskemét house price index in period t. $\tilde{Y}_{1t}$ denotes the counterfactual state, or the house price index value which would have been observed in Kecskemét in the absence of a treatment.

As we cannot directly assess how house prices in Kecskemét would have changed if the Mercedes factory had not been constructed, the extent of the treatment's impact can be estimated by approximating the counterfactual state with a synthetic control: $\hat{\beta}_{1t} = Y_{1t} - \hat{Y}_{1t}$.

The estimation of the counterfactual state uses information on the donor pool units. The analysis shows the units, which are the Hungarian county seats, with index *j*. The treated unit under review is denoted by index *1*. The donor pool contains J units that sufficiently resemble the unit under review, but are nonetheless untreated; therefore, $j = 1, \dots, J+1$. The periods, which are quarters in the study, are shown with an index $t = 1, \dots, T$, and the treatment, which is the factory investment, appears in period $T_0 + 1$. Thus, T_0 shows the last month of the pre-treatment period.

The counterfactual state is approximated by generating a synthetic control, which is a linear combination of the donor pool units, $\hat{Y}_{1t}(W) = \sum_{j=2}^{J+1} w_j Y_{jt}$, where w weights are non-negative and their sum is 1, and can be described by the vector $W = (w_2, \dots, w_{J+1})'$.

[12] Some of these concerns are related to the case when several units are affected (see, for example, Abadie–L'Hour 2021, Dube–Zipperer 2015). Other studies developed a regression-based estimation process to produce synthetic controls, where the weights of the donor pool can be negative and their sum is not limited to one (e.g. Chernozhukov et al. 2021, Doudchenko–Imbens 2016). Finally, certain studies focus on the estimation bias and its correction (Abadie–L'Hour 2021, Ben-Michael et al. 2021).

Ideally, weights should be determined with a minimum approximation error; that is, a minimal difference between the synthetic control and counterfactual state in the post-treatment period: $\tilde{Y}_{1t} - \hat{Y}_{1t}(W)$. However, the counterfactual state is unobservable. Therefore, Abadie–Gardeazabal (2003) and Abadie et al. (2010) suggested determining the weights by minimising the difference between the actual and synthetic control time series (for the outcome variable and other predictors) in the pre-treatment period.

The estimated impact is derived from the difference between the actual time series' of the treated unit and synthetic control, which can be written as follows in a given period: $\hat{\beta}_{1t} = Y_{1t} - \sum_{j=2}^{J+1} w_j Y_{jt}$. A relevant question in many impact assessments is the average impact of a treatment in a given period. The average treatment effect on the treated unit (ATT) in the post-treatment period is:

$$ATT\ (T_0 + 1, T) = \frac{1}{T-T_0-1} \sum_{t=T_0+1}^{T} \beta_{1t} \qquad (1)$$

To produce a reliable estimate of the extent of the treatment's impact using the SCM, the synthetic control should fit the actual time series of the treated unit well during the pre-treatment period. Ideally, the fit should be good, while considering several economic predictors besides the outcome variable. The predictors usually include the values of the outcome variable from *m* pre-treatment periods and *r* variables describing the economic environment that has forecasting power for the outcome variable; thus, $K = r + m$.

Let X_1 denote the $Kx1$ matrix of predictors for the treated unit and X_0 denote the KxJ matrix of the predictors for the donor pool units. Synthetic control is established using an optimisation process, where first an inner and then an outer optimisation task must be solved.

The inner optimisation task is used to determine the donor weights to ensure that the actual time series are as close to the synthetic control time series as possible in terms of the outcome variable and predictors in the pre-treatment period; accordingly, a convex combination of the columns of X_0 is sought that best approximates X_1. A W weight vector is sought with a given V matrix, where the units are non-negative and their sum is 1, and the $\|X_1 - X_0\, W\| = \sqrt{(X_1 - X_0\, W)'\, V\, (X_1 - X_0\, W)}$ expression is minimised, where V is the non-negative diagonal matrix of the KxK predictor weights and W denotes the weight vector that determines the role of the control units in producing the synthetic control.

Next, the aim of the outer optimisation task is to determine the units in the V matrix. Typically, no perfect fit can be achieved for all predictors (Klößner–Pfeifer 2018). Thus, the $V = (v_1, \dots, v_K)$ weight vector assigns different importance to the different predictors (v_k is the weight of the kth predictor). Then, a better fit is expected for predictors with a greater weight. However, determining the predictor weights is a non-trivial task. Usually, the aim is to assign greater weights to predictors that have a better forecasting capacity for the outcome variable.

A data-driven approach is used to determine the predictor weights in line with the proposal of Abadie–Gardeazabal (2003), and Abadie et al. (2010).[13] The diagonal positive definite matrix is used for the predictor weights, where the mean squared prediction error (MSPE) value of the outcome variable is the lowest in the pre-treatment period. Therefore, $(Z_1 - Z_0\, W^*(V))'\,(Z_1 - Z_0\, W^*(V))$ must be minimised, where Z_1 denotes the $Mx1$ vector containing the Y values of the treated unit in the pre-treatment period, and Z_0 shows the MxJ matrix containing the Y values of the donor pool units in the pre-treatment period.

Here, a generalised and expanded version of the standard SCM, or the MSCMT is used due to the two features detailed in the Appendix (3. The multivariate synthetic control method estimator).

Methodological of drawing inference

Conclusions can be drawn from the estimates when the pre-treatment fit is appropriate. Therefore, a sufficiently long pre-intervention period is needed because the effect estimated using the SCM can be considered to be reliable if the actual time series in the pre-intervention period is well approximated by the synthetic control time series.

Several indicators can be used for assessing the goodness of the pre-treatment period fit, which typically have similar underlying logic. We use both Ferman's $\tilde{R}^2$ (Ferman et al. 2020) and modified Cohen's D indicators (Hollingsworth–Wing 2020) in our analysis.

Ferman's $\tilde{R}^2$, or the pre-treatment normalised mean squared error index, can be expressed as follows:

$$\tilde{R}^2 = 1 - \frac{\sum_{t=1}^{T_0}\left(Y_{1t} - \hat{Y}_{1t}^{N}\right)^2}{\sum_{t=1}^{T_0}\left(Y_{1t} - \frac{\sum_{t=1}^{T_0} Y_{1t}}{T_0}\right)^2} \qquad (2)$$

The closer the value of the indicator is to 1, the better the fit of the synthetic control time series in the pre-treatment period. $\tilde{R}^2 = 1$ implies a perfect fit. Ferman et al. (2020) examined the goodness of fit for two cut-off values, 0.8 and 0.95.

The formula for the modified version of Cohen's D indicator, frequently used in the matching literature, is as follows:

[13] Several studies use a cross-validation technique to determine predictor weights (e.g. Abadie et al. 2015, Becker et al. 2018, Donohue et al. 2019, Kellogg et al. 2021, Klößner–Pfeifer 2018, Xu 2017); however, it should not be used for all databases. The pre-treatment period must be divided into training and validation periods during cross-validation. The validation sample should be at least as long as the post-treatment period (Hollingsworth–Wing 2020). The present impact assessment contains 28 pre- and 28 post-treatment periods each. Thus, the validation sample would exhaust the entire pre-treatment period and no observation would remain for the training sample, whose duration is usually set to be longer than the validation period. Therefore, the database available for the impact assessment is not considered to be suitable for the application of the cross-validation method (e.g. Becker et al. 2021), and the examination determines predictor weights in a data-driven manner.

$$D_j = \frac{1}{T_0} \sum_{t=1}^{T_0} \left| \frac{Y_{jt} - \hat{Y}_{jt}^N}{\sqrt{\frac{1}{T_0} \sum_{t=1}^{T_0} \left(Y_{jt} - \bar{Y}_j\right)^2}} \right| \tag{3}$$

An advantage of this indicator is that it is independent of the units of measure; thus, even values calculated for different indicators can be compared.

There are several factors during comparative impact assessments that make it difficult to draw statistical inferences. Abadie et al. (2015) emphasised the small sample size and lack of randomisation, as well as the fact that donor units are not determined by probabilistic sampling. In this study, the intervention only affects a single unit, Kecskemét; therefore, no sampling error is present, and no standard hypothesis testing can be used.

If the number of control units is limited, large-sample size methods must be replaced with alternative solutions to determine whether the estimated effect can be considered significant. Abadie et al. (2010) proposed a process similar to permutation tests by conducting spatial and temporal placebo tests. In a spatial placebo test, randomness concerns the unit affected by the intervention; here, this is, the county seat where the car factory is built. Meanwhile, during temporal placebo testing, the randomness of the intervention's timing is tested.

Spatial placebo tests are conducted to establish whether the estimated effect for the unit affected by the intervention is substantially greater than the placebo effects estimated for the donor pool units. Here, we assess how often an effect of a similar magnitude would be achieved if the county seat where the intervention is administered is chosen randomly. Specifically, one of the donor pool units is assumed to be a treated unit and the actually treated unit is added to the donor pool before producing the synthetic time series and estimating the pseudo-effect. This is performed for all donor pool units. Subsequently, the distribution of the pseudo-effects estimated for the donor units is compared with the actual effect for the treated unit. The impact of factory construction on house prices can be considered significant if the estimated impact for the actual treated unit is distinct from the distribution of placebo effects.

To determine the statistical significance of the estimated effect, the ranked *p* value is calculated when assessing the placebo testing results. The distribution of the pseudo-effects derived from the placebo tests forms the basis for comparison. The ranking of the actual treated unit can be determined using several indicators. If the indicator calculated for the jth unit is denoted as r_j, the *p*-value is calculated using the following formula:

$$p = \frac{\sum_{j=1}^{J+1} I_+(r_j \geq r_1)}{J+1} \tag{4}$$

where $I_+(\cdot)$ is an indicator function, which equals 1 if the expression in the parentheses is satisfied. Here, the number of units in the donor pool is 15. Therefore, the lowest possible *p*-value in permutation testing is 6.25%.

In spatial placebo tests, donor pool units are sometimes excluded if the synthetic control has a much weaker fit in the pre-treatment period than the unit under review. There is no clear guideline on which control units should be excluded when assessing placebo tests. For example, Abadie et al. (2010) excluded donor pool units where the pre-intervention MSPE is more than twenty, five, or two times larger than that for the unit under review. The results are typically generated with various values because there is no objectively ideal cutoff value for excluding donors. Therefore, we use two indicators that utilise the entire donor pool, and consider and manage scenarios in which the synthetic control for any donor pool unit has a relatively poor fit.

The ratio of the RMSPE calculated for post- and pre-intervention periods is a widely used indicator for assessing the significance of the estimated impact (see Abadie et al. 2010, 2015, Klößner–Pfeifer 2018). The RMSPE value of the synthetic control, which measures the fit of the synthetic control time series to the actual time series, is written as follows if $j = \{1, \ldots, J+1\}$ and $0 \leq t_1 \leq t_2 \leq T$:

$$RMSPE_j\,(t_1, t_2) = \left(\frac{1}{t_2 - t_1 + 1} \sum_{t=t_1}^{t_2} \left(Y_{jt} - \hat{Y}_{jt}^N\right)^2\right)^{\frac{1}{2}}, \quad 0 \leq t_1 \leq t_2 < T, \quad j = \{1, \ldots, J+1\} \tag{5}$$

where $\hat{Y}_{jt}^N$ is the synthetic control of the outcome variable in the *t*th period if the jth unit is considered treated and the other units form the donor pool in the given placebo test. The distribution of the ratios of post- to pre-treatment RMSPE (PPR_j) is calculated for the donor pool and treated units as follows:

$$PPR_j = \frac{RMSPE_j(T_0+1, T)}{RMSPE_j(1, T_0)} \tag{6}$$

The rationale behind this ratio is that a large impact in the post-treatment period cannot be considered significant if the difference between the actual and synthetic time series is also large in the pre-treatment period. Thus, from an impact assessment perspective, the effect estimated for Kecskemét can be considered significant if the ratio of the post- to the pre-treatment RMSPE is large for the city and exceeds the value estimated for donors.

The ratio of post- to pre-treatment RMSPE assesses the significance of the estimated effect for the entire post-intervention period; however, another aspect worth exploring is the time profile of the estimated effect. For this, similar to Klößner–Pfeifer's (2018) exercise, we compute a relative gap time series. The pseudo-gaps yielded by the placebo test are adjusted depending on the fit of the synthetic control time series in the pre-treatment period instead of narrowing down the donor pool. The values of the relative-gap time series are described by the following:

$$RG_{jt} = \frac{Y_{jt} - \hat{Y}_{jt}^N}{RMSPE_j(1, T_0)}, \quad T_0 + 1 \leq t \leq T \tag{7}$$

The empirical significance levels corresponding for *PPR* and relative gap (RG) are calculated as follows:

$$p = \frac{\sum_{j=1}^{J+1} I_+(PPR_j \geq PPR_1)}{J+1} \;;\quad p_t = \frac{\sum_{j=1}^{J+1} I_+(RG_{jt} \geq RG_{1t})}{J+1} \tag{8}$$

where $I_+(\cdot)$ is an indicator function.

Results

Model specification

Units with a similar intervention on the analysis time horizon should be excluded from potential controls (Abadie et al. 2015). As Győr experienced several major automotive industry plant expansions related to the Audi factory during the study period coupled with a dynamic expansion of employee numbers, it is excluded from the potential controls. In addition, settlements whose features differ substantially from the city where the intervention under review has been administered should also be excluded from the sample. Salgótarján is excluded because of its exceptionally low level of development relative to other county seats. Further, the time horizon of the estimation needs to be reduced by two years if this county seat is included. Finally, the house price index estimated for Salgótarján does not seem reliable due to the low number of transactions.

Next, the database used for the calculations has been narrowed in terms of its time horizon. This is because the house price indices for county seats and nearby settlements, which serve as outcome variables, can only be generated for the early 2000s with some uncertainty owing to the available data. Therefore, the estimation has been carried out on a smaller panel sample of 16 county seats: Kecskemét, where the factory construction under review has occurred, and 15 potential control cities. This narrowed database includes values for 56 quarters between 2004 and 2017.

The time of the treatment, or the Mercedes factory's investment (T_0), is set as 2010 Q4. This makes it the last quarter of the pre-treatment period, which is followed by the post-treatment period. Factory construction in Kecskemét was announced in 2008. Although production only started in 2012, we argue that because of the factory's trial run that began in 2011, gradual increase in the workforce, and housing demand of future employees that appeared prior to their employment, the additional housing market demand may have exerted an impact in the year before production started. Essentially, the pre-treatment period covers 2004 Q1–2010 Q4, while the post-treatment period covers 2011 Q1–2017 Q4, containing 28 quarters each.

The synthetic control time series is represented by the [15X1] $W = (w_1, w_2, \dots, w_{15})$ vector, where the sum of the weights is $w_1 + w_2 + \cdots + w_{15} = 1$. The W weight vector is determined by minimising the differences between the features of the treated and control cities during the pre-intervention period. Characteristics that have greater forecasting power for the outcome variable have a larger weight.

Using the variables presented in the second section, 14 predictors, together with the outcome variable, are included in the estimation. The inclusion of the outcome variable among predictors is a widely accepted practice, although individual studies differ in terms of the outcome variable values from which periods are used as predictors. Ferman et al. (2020) theoretically examined the issue of specification searching with respect to the SCM using Monte Carlo and actual data simulations. Based on their results, they formulated recommendations for including the outcome variable as a predictor to limit specification searching. Accordingly, the impact assessment uses the values of the outcome variable from the first half of the pre-treatment period among the predictors. This is intended to ensure that the synthetic control group includes cities that resemble Kecskemét in terms of housing price trends prior to the investment.

Baseline model

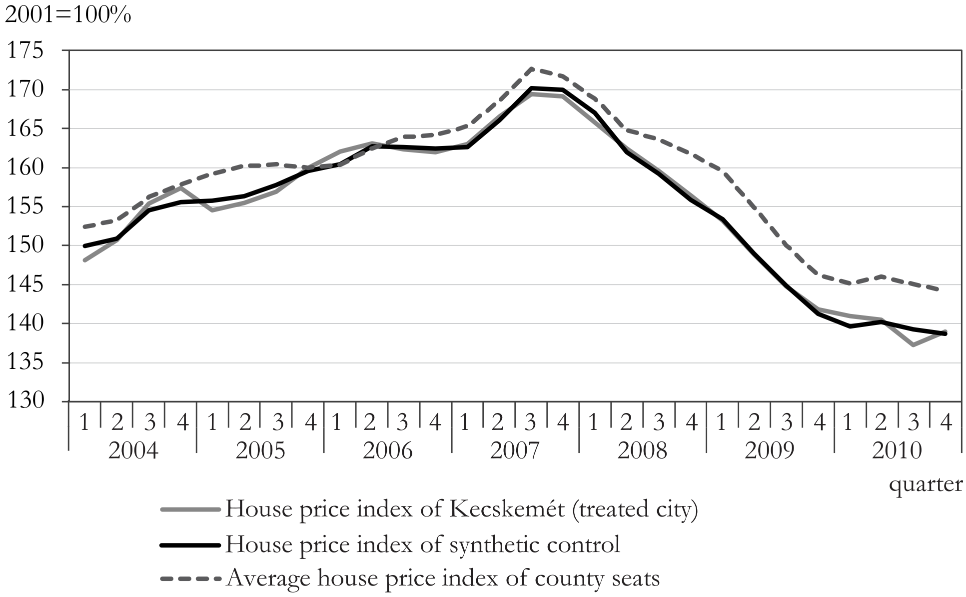
The baseline model is chosen as follows: we run all model combinations that include 2–6 predictors, yielding 6,461 model estimates[14]. The number of predictors is not increased above six because in the overwhelming majority of the estimated models, no six covariates received significant weighting simultaneously. The baseline model is selected based on the best fit for the pre-intervention period (lowest RMSPE).

The baseline model uses the following six predictors: house price index, corporate equity per capita, number of hospital beds per 1,000 people, average square metre prices, ratio of community housing to the stock of dwellings, and shortest travel time to Budapest by car. The first two predictors are assigned a major weight; thus, the synthetic Kecskemét house price index is generated as a convex combination of house prices in the county seats in the donor pool which most closely resemble Kecskemét in terms of house price trends and the presence of the corporate sector. Szolnok has the highest weight (28.2%), followed by six cities with a weighting of 9% each, and another with 0.4%; accordingly, the synthetic control time series comprises the weighting of eight cities. Figure 4 shows how much better the house price index of the synthetic control 'city' estimated with the baseline model fits onto the development of house prices in Kecskemét than if a simple average of all the county seats in the donor pool is used.

[14] The database was compiled and statistical calculations were conducted with the R software, Version No. 3.5.1. The synthetic control estimates and the results were produced using the MSCMT package.

Figure 4

House price index of Kecskemét, the baseline model and average house price index of the donor cities in the pre-treatment period



The results presented below are for the baseline model with the lowest RMSPE and the set of models with up to 5% higher RMSPE values; the number of latter models is 108. Figure 5 shows that these models, which still have a good fit and yield a synthetic control time series close to the baseline model. Accordingly, similar results can be achieved using various combinations of predictors. Overall, the synthetic control time series fits well onto the Kecskemét house price index in the pre-intervention period. Therefore, the synthetic control time series seems to be a good approximation of the unobservable counterfactual state, or how house prices in Kecskemét would have evolved in the absence of the Mercedes project. In the pre-intervention period, the fit of the synthetic control is examined using two indicators: Ferman's $\tilde{R}^2$ and Cohen's D. In the former case, the closer the value is to 1, the better the fit. The fit of the models is appropriate even with a stricter cutoff value of 0.95. The baseline model has an value of 0.99. With respect to Cohen's D, an indicator lower than 0.25 and falling as close to 0 as possible is expected; in the baseline model, it is 0.0086. All 108 models yield similar values. Thus, both indicators under review show that the fit of the models is appropriate and they can be used to draw conclusions from the results.

Figure 5

Actual time series of the house price index in Kecskemét, and the minimum, maximum, and median of the baseline model and models with up to 5% higher RMSPE

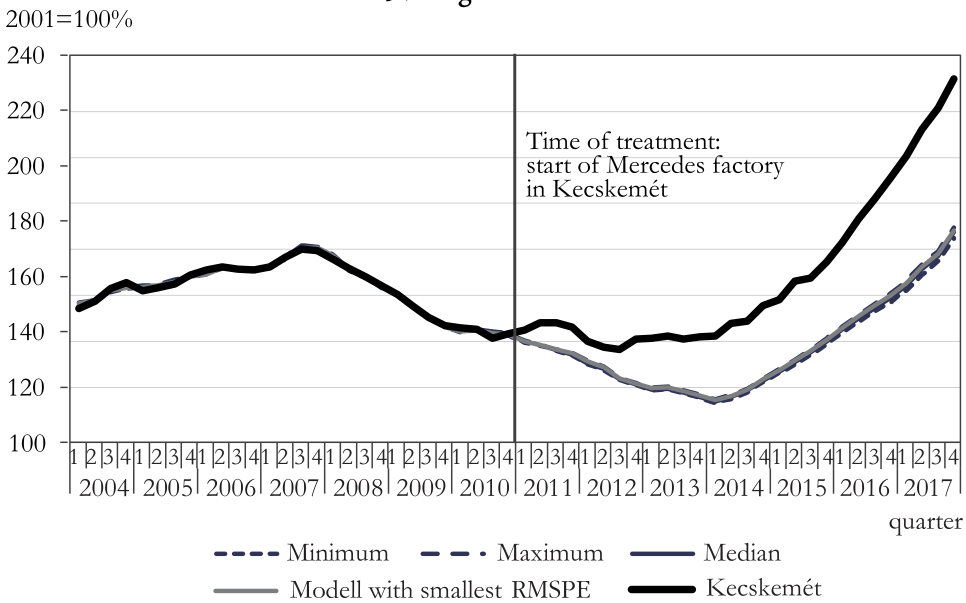
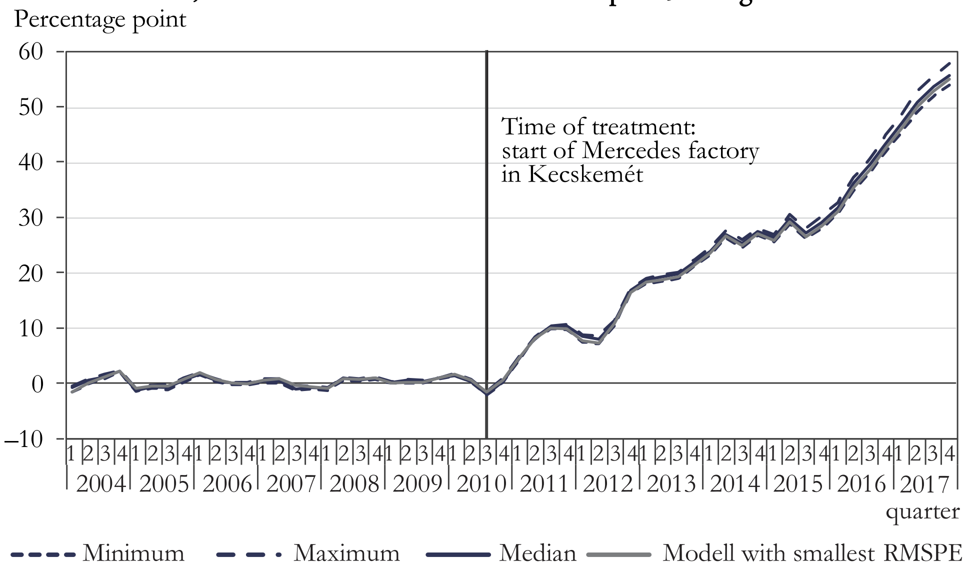


Figure 6

Actual time series of the house price index in Kecskemét, and the gap between the synthetic controls in the baseline model, as well as the minimum, maximum, and median of the models with up to 5% higher RMSPE



According to the baseline model results, house prices in Kecskemét and nearby settlements increased by 39.2 percentage points more between 2010 Q4 and 2017 Q4 than they would have without the Mercedes factory, and its economic and social spillover effects. The estimated effect attributes an additional housing price appreciation of 15.1 percentage points to the investment project in the first three years. This impact is being observed because besides the minor fluctuations, house prices in Kecskemét did not drop significantly in the three years after 2010, whereas the estimates show that they would have decreased without the Mercedes project. Similar to the difference in housing price changes with an expanding time horizon, the results increasingly reflect not only the effects directly attributable to the project but also its spillover effects. The Mercedes project examined here led to the appearance of many suppliers and related services, creating more work opportunities. In 2011–2012, Kecskemét had the second-highest balance of net domestic migration among the county seats; thus, most county seats had negative net migration, while Kecskemét had a positive balance of people moving there and leaving. After the factory was setup, the after-tax income per capita also increased more than in other county seats, indicating that newly created jobs improved the purchasing power of the population. These factors may have contributed to faster development of the city and sustained higher housing market demand. As we show in the second section, the supply of new housing responded in a timely and positive manner to the growing housing market demand. Thus, even before the investment was implemented, the number of new housing constructions in Kecskemét was better than in the other county seats.

Examining the 108 models with an RMSPE up to 5% higher than the model with the lowest RMSPE shows that the range of the effects of factory construction on house prices is quite narrow. The synthetic control time series describing the smallest house price appreciation yields an additional impact of 41.1 percentage points, whereas that describing the largest increase yields an additional impact of 38.9 percentage points (Table 1). Overall, the construction of the Mercedes factory in Kecskemét substantially affected house prices in the city and neighbouring settlements. Moreover, the difference in house prices continued to grow in the years following the factory's construction.

According to Abadie et al. (2015), the estimates may be distorted if the intervention has a spillover effect on the donors in the synthetic control, which refers to other county seats in the present analysis. This issue probably does not arise here because house prices are not territory specific. Furthermore, we can assume that the Mercedes project in Kecskemét (and increased property prices) did not significantly affect property prices in other county seats. County seats outside Budapest are far from each other. Thus, an investment project in Kecskemét is unlikely to attract commuting workers from other county seats.

Table 1

Differences between Kecskemét and estimated synthetic controls in the change in house prices relative to 2010 Q4 (based on the distribution of 108 models)

(percentage points)

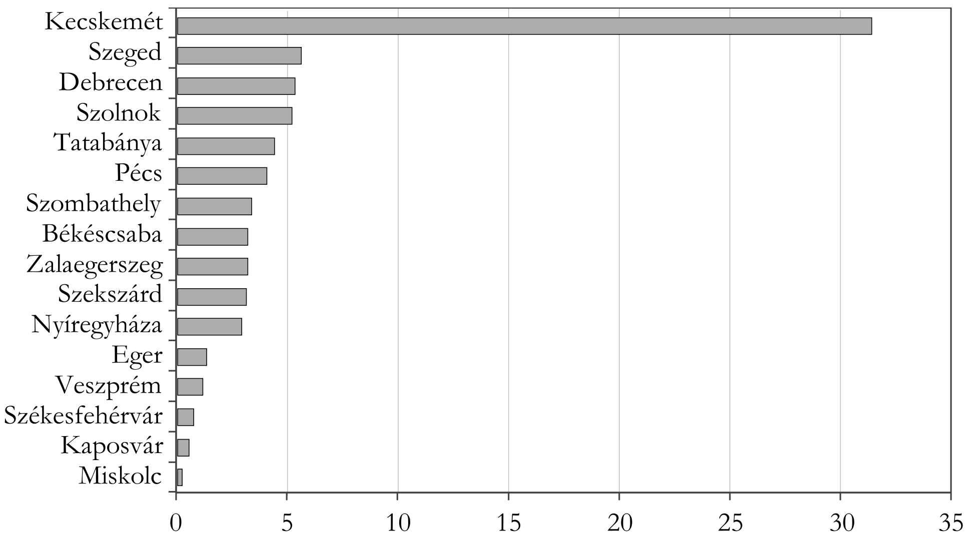
Year	Minimum	Median	Maximum	Smallest RMSPE model
2013	15.6	15.7	15.1	15.1
2014	19.6	19.7	19.2	19.1
2015	21.2	20.8	20.0	20.1
2016	31.7	31.1	30.1	30.2
2017	41.1	40.0	38.9	39.2

Inference about the model

We test the significance of the results using two ways. First, we run spatial placebo tests using the baseline model, assuming that instead of the city affected by the investment project, or Kecskemét, another county seat in the donor pool is treated. Then, the synthetic control for that city is calculated by adding Kecskemét to the donor pool. Based on the synthetic control derived in this manner, we compute the gap from the actual time series, or the pseudo-effect. Subsequently, this exercise is repeated for all other county seats in the donor pool and two indicators are calculated in all placebo runs: the ratio of post- to pre-treatment RMSPE, and relative gap.

Figure 7

Ratio of post- and pre-treatment RMSPE values in Kecskemét and donor county seats

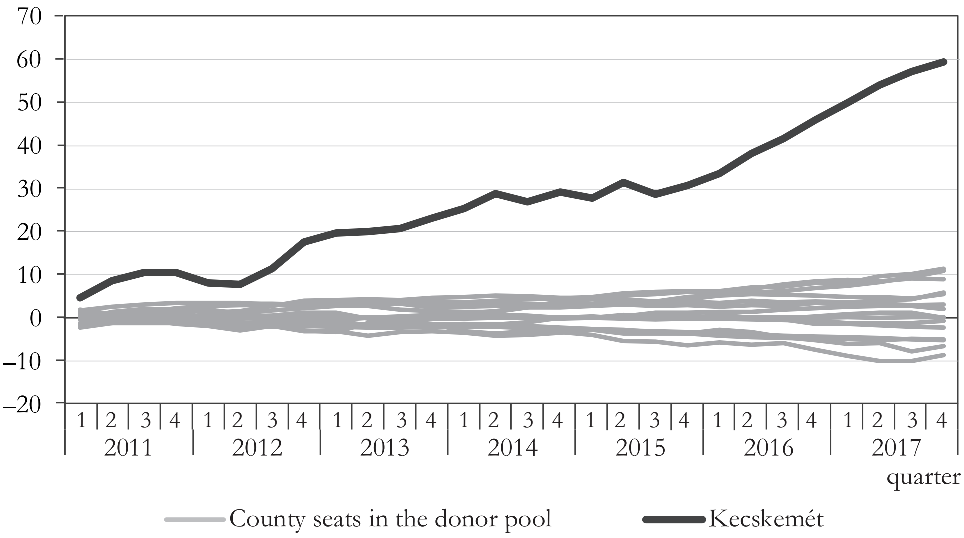


Note: The pre-treatment period covers 2004 Q1–2010 Q4, while the post-treatment period covers 2011 Q1–2017 Q4.

The originally estimated effect can be considered significant if in the actual treated city, which is Kecskemét, the post-intervention RMSPE value of the synthetic control is large relative to the pre-intervention RMSPE, and exceeds the value for the county seats in the donor pool. Figure 7 shows that Kecskemét stands out from the remaining county seats. The post-intervention RMSPE is 31 times higher than the pre-intervention RMSPE. The highest ratio of post- to pre-treatment RMSPE estimated for Kecskemét indicates that the probability of estimating a larger impact during the random permutation of the treated status than in Kecskemét is 6.25%; thus, the null hypothesis that the intervention had no effect can be rejected at a standard 10% significance level.

Figure 8

Relative gap time series for Kecskemét and county seats in the donor pool



The ratio of post- to pre-treatment RMSPE is used to measure the significance of the effects over the entire post-treatment period. Furthermore, the results of the spatial placebo test are employed to examine the entire time profile and significance of the estimated effect. The time series of the pseudo-effects derived for the county seats in the donor pool are normalised with the pre-intervention period RMSPE values. This is because the large difference measured in the post-intervention period only indicates an appropriate estimated impact if the pre-intervention period has a good fit. Hereafter, these are referred to as the relative gap. In Figure 8, the dark line shows the relative gap estimated for Kecskemét, whereas the grey lines denote the relative difference in the synthetic control time series estimated for the county seats in the donor pool. The Kecskemét time series stands out, as the relative gap for Kecskemét is much larger than in any other county seat. This suggests that the Mercedes project had a significant role in the fact that house prices in Kecskemét increased faster than those in other Hungarian county seats. Over the entire post-intervention period, the relative gap is the largest for Kecskemét; therefore, the empirical significance level is 6.25% for all periods.

Strengths and limitations

We argue that the SCM is suitable for examining the research question. The impact assessment presented here correspond to the conditions mentioned by Abadie (2021). First, we can assume that stable unit treatment value assumption (SUTVA) is satisfied (see in the Appendix [2. Indentifying assumptions]). Thus, the outcome of the units under review is independent of whether the other units received treatment[15]. No major spillover effects are assumed across the housing markets of Hungarian county seats, and the housing market of Kecskemét is sufficiently independent of that of the other county seats under review. Further, the observed effect is large enough and outcome variable of the treated unit is not too volatile. Appropriate donor units are available to produce synthetic controls. Both the pre- (28 periods) and post-intervention samples (28 periods) are of sufficient size. We consider the potential anticipation effect when determining the timing of treatment. We believe that the endogeneity issue (Abadie–Gardeazabal 2003) is not relevant in the present case because the change in house prices is unlikely to have influenced Mercedes' decision to choose a city for its factory. In other words, the choice of the treatment unit is not influenced by the variables under review.

The model specification can substantially influence the results of the impact assessment via two factors: the set of predictors and choice of donor units (Abadie 2021). Therefore, robustness analysis should be performed using these factors. Thus, we perform synthetic control estimations for various combinations of predictors. The results are detailed in fourth section, and the robustness analysis of the donor pool is presented in fifth section. The robustness analyses confirm the reliability of the results.

Finally, this study has several limitations. The pre-treatment period sample is certainly not small, as many studies have performed impact assessments with the SCM using fewer observations (e.g. Baccini et al. 2014, Bohn et al. 2014, Dustmann et al. 2017, Hinrichs 2012, Kleven et al. 2013, Lindo–Packam 2017). Nonetheless, because of the features of the database used in this study, performing cross-validation to

[15] During the sample period, there was one event similar to our treatment: Audi's factory expansion in Győr. This raises a concern because if i) the Audi factory expansion had significant spillover effects to nearby county seats which are part of the donor pool, and ii) the synthetic control is (at least partly) constructed from these nearby county seats, then the estimated effect could be biased. This concern was alleviated by the results of our exercise, whereby we used the Audi factory expansion as a treatment and the three nearby county seats (Veszprém, Szombathely, and Tatabánya) as treated cities one by one. Both the ratio of post- to pre-treatment RMSPE, and the relative gaps showed that the estimated effects are statistically not significant. Thus, the Audi factory expansion does not influence our results related to the Mercedes factory investment in Kecskemét.

determine the predictor weights is not worthwhile. Furthermore, the estimated effect may not be entirely attributable to the Mercedes project; it may also capture the impact of other investments on Kecskemét. However, the significance of the Mercedes factory investment is supported by the fact that the factory's arrival was the main driver behind infrastructure development in the city, such as motorway development (construction of a bypass to shorten the distance to the capital), railway infrastructure development that included the creation of industrial sidings, community housing development, development of a university department and nursery school, and kindergarten construction.

Robustness analyses

During the estimations, several decisions have been made concerning the methodology that may influence the results; therefore, a robustness analysis of the model must be conducted from various perspectives. Specifically, the baseline model with the lowest RMSPE is contrasted with the features of the model with the smallest RMSPE from various runs. The main results are presented in Table 2, and further details can be found in Table A3 of the Appendix.

First, the estimation is carried out using all pre-treatment period values of the house price index. Several applications of the SCM (e.g. Billmeier–Nannicini 2013, Bohn et al. 2014, Liu 2015, Stearns 2015) include all pre-intervention period values of the outcome variable among the predictors. In this case, the other predictors do not play a role in producing the synthetic control, as noted by Kaul et al. (2021), both theoretically and empirically. According to Botosaru–Ferman (2019), the bias of the synthetic control estimation procedure may be limited even if only the pre-intervention period values of the outcome variable are used as predictors, resulting in a synthetic control with good fit. In our case, the latter holds true because the fit of the synthetic control is good and similar to the baseline model: the value of $\tilde{R}^2$ is 99% and Cohen's D is 0.0086. The weights of the donor cities are similar to those of the baseline model, and the estimated impact only differs from the baseline model's effect by one percentage point until 2017.

Table 2

Main results of the models' robustness analysis

Denomination	Baseline modell	Robustness analysis: Without economic predictors, with all the pre-treatment period values of the outcome variable	Robustness analysis: Shorter sample, pre-intervention period starts in 2005 Q1	Robustness analysis: Shorter sample, pre-intervention period starts in 2006 Q1	Robustness analysis: Excluding the donor with the largest weight, i.e. Szolnok
Estimated effect untill 2017 (percentage points)	39.22	38.05	44.26	32.92	33.34
PP-ratio, Kecskemét	31.41	31.63	38.12	35.21	24.60
PP-ratio, maximum without Kecskemét	5.60	6.00	7.13	9.45	6.40
R squared (á la Ferman et al., 2020)	0.990	0.991	0.993	0.950	0.987
modified Cohen's D (á la Hollingsworth and Wing, 2020)	0.009	0.008	0.007	0.004	0.010

Denomination	Robustness analysis: Excluding the donor with the second largest weight, i.e. Eger	Robustness analysis: Seed=50 for the random number generator	Robustness analysis: Genetic algorithm is used for outer-optimization	Robustness analysis: Testing anticipation effect: pre-intervention period ends in 2011 Q4	Robustness analysis: In-time placebo test: pre-intervention period ends in 2017 Q4
Estimated effect untill 2017 (percentage points)	33.68	39.22	39.34	38.02	2.23
PP-ratio, Kecskemét	25.49	31.41	31.42	11.60	0.00
PP-ratio, maximum without Kecskemét	5.32	5.60	5.60	4.68	0.00
R squared (á la Ferman et al., 2020)	0.988	0.990	0.990	0.943	0.982
modified Cohen's D (á la Hollingsworth and Wing, 2020)	0.010	0.009	0.009	0.019	0.017

Note: The results in the table refer to the model with the smallest RMSPE value during the pre-intervention period.

The model is also examined using a shorter sample by first removing one year and then two years from the beginning of the pre-intervention period. The R^2 is somewhat lower at 95%, for the sample which is two years shorter and starts in 2006; still, it exhibits a good fit. Except Szolnok, similar donor cities are assigned large weights. The estimated effect decreases slightly from 39 to 33 percentage points.

Robustness analysis also cover various technical parameters of the model. First, during the calculation, the method for generating random numbers is changed. Second, instead of using the DEopC algorithm to solve the outer optimisation task, an alternative heuristic optimisation procedure, the genetic algorithm, is used in line with the recommendation of Becker–Klößner (2018). These changes modify the estimated effect by only a few tenths, and the weights of the donor pool cities and predictors do not shift significantly.

Next, we examine whether the results are sensitive to the exclusion of the main donor cities from the estimation. The baseline model is re-estimated by first excluding Szolnok, which has the highest weight, and then Eger, which has the second highest weight. The results show that the model fit does not change, as the R^2 is unchanged at 99%. The estimated effect is 33–34 percentage points, which is somewhat lower than the baseline model's 39-percentage point value. Therefore, the results are robust if the two donors with the highest weights are successively excluded from the estimation.

In line with the proposal by Abadie (2021), a backdating exercise is performed to examine the credibility of the synthetic control. Specifically, the intervention is moved to an earlier time during the estimation period (in-time placebo test). We believe that the validity of the synthetic control estimation will be confirmed if the difference corresponding to the actual pre-intervention period is around zero. This is because this demonstrates that the synthetic control accurately reflects the development of house prices in the treated city prior to the intervention. Accordingly, the time of the intervention is set as 2007 Q4 instead of 2010 Q4. The results show that until 2010, the time of the actual intervention, the estimated effect is minimal and only exhibits a 2 percentage point difference between the house price indices in the synthetic control and Kecskemét. Moreover, the extent of the impact is not significant, as Kecskemét has an average *PPR*, and its relative gap time series also falls short of those of several other donor cities. The reliability of the synthetic control and effect estimated using it will improve if the effect is significant after the actual intervention time, even if the intervention time is set much earlier than the actual date (Abadie et al. 2015). When the treatment date is set as 2007 Q4, the estimated effect until late 2017 is 44.5 percentage points, which is of a similar magnitude to the baseline model effect of 39 percentage points; therefore, the backdating exercise confirms the results of the impact assessment.

We also examine what would happen if the intervention date is pushed slightly forward to 2011 Q4. Again, the estimated effect is approximately similar to that of the baseline model; therefore, in Kecskemét, the ratio of post- to pre-treatment RMSPE is much lower than in the baseline model, but still stands out among the remaining cities.

Overall, the results of the robustness analyses show that the spatial placebo tests yield a much higher ratio of post- to pre-treatment RMSPE for Kecskemét than for the other cities in all cases. Furthermore, Kecskemét fares best in terms of the distribution of relative gaps. Thus, the results hold even after using the various robustness analyses. Furthermore, the effects estimated by the different model runs are also dispersed between 33 and 42 percentage points, which is around the baseline result.

Conclusions

This study assessed the impact of Mercedes' investment in Kecskemét on housing prices in the city and its neighbouring settlements. The extent of this effect was determined by using a SCM. The estimations employed the MSCMT method, which is a generalised and expanded version of the standard SCM.

The estimation was performed on a panel sample containing values from 56 quarters for 16 Hungarian county seats. The pre-treatment period covered 2004 Q1–2010 Q4, while the post-treatment period covered 2011 Q1–2017 Q4. The estimated synthetic control fit well with the actual data in the pre-intervention period; therefore, the model was appropriate.

The baseline model results show that following the completion of the Mercedes factory, house prices in Kecskemét increased by 39.2 percentage points more between 2010 Q4 and 2017 Q4 than they would have without the investment project. The significance of the results was tested in two ways: the ratio of post- to pre-treatment RMSPE was used to assess the significance of the impact estimated for the entire post-treatment period, and the relative gap time series was used to examine the time profile and significance of the estimated effect. The significance tests suggest that the Mercedes project played a significant role in the fact that house prices in Kecskemét increased faster than in other Hungarian county seats. The robustness of the results was analysed from various perspectives because there are several decision points during SCM modelling that could affect the estimates. Robustness analyses confirmed the results, and the estimated impact was significant in all cases.

The results show that in an economy in which the automotive industry is a dominant sector, a large-scale investment similar to the Mercedes project significantly affects certain segments of the local economy, such as the property market. Therefore, the methodological implementation and main conclusions of this study can provide useful information to decision-makers both during planning and after implementation to assess the subsequent effects of an investment on a region's broader economic and social environment.

Appendix

1. Advantages of the SCM

The synthetic control method (SCM) is used because it can be particularly useful when a single unit is treated, such as Kecskemét in our case, and relatively few similar control units are available, such as the county seats. With such data, it is rare to have a single control unit that has several features which resemble the treated unit well enough to allow reliable conclusions to be drawn by comparing these units. The pre-intervention features of the treated unit can be more accurately approximated with some combination of donor units than with a single untreated unit (Abadie et al. 2015). The same holds true here, as we have no county seats where the parallel trend assumption holds true with respect to Kecskemét in the pre-treatment period (Figure 1); thus, no county seat by itself would approximate the counterfactual state.

Another advantage of the method is that the control group includes units that can be considered similar to the treated unit based on the features relevant for the outcome variable. The principle is similar to propensity score matching (Rosenbaum–Rubin 1983, 1985), which cannot be used in a sample with only one treated unit.

Moreover, with such a data structure, traditional regression methods are typically unsuitable as they can be most useful in an impact assessment with a large number of observable units and repeated occurrence of an event/intervention (Abadie 2021). A further advantage of the SCM compared to regression estimates is that it transparently shows the weights with which control units contribute to producing the counterfactual time series. Further, it shows the extent of which the synthetic and treated time series resemble each other in terms of the outcome variable and predictors. Moreover, there is no extrapolation as the weights are positive and their sum is exactly 1 (Abadie et al. 2010).

In the case of comparative impact assessments, the small sample size in itself does not prevent quantitative conclusions. The main obstacle is rather the lack of an explicit mechanism for picking the control units (Abadie et al. 2015). This is partly why the SCM offers a great advantage as the counterfactual state of the treated unit is not approximated by the time series of an ad-hoc control unit. Instead, a synthetic time series is generated in a transparent, data-driven process. This allows one to run placebo tests, which can then be used to draw quantitative conclusions with the SCM (Abadie et al. 2010, Abadie 2021).

2. Identifying assumptions

Two assumptions are usually made in using treatment evaluation methods (Ferracci et al. 2014), which provide a basis for considering the estimate as the result of a causal relationship.

The *conditional independence assumption* (*CIA*) states that after conditioning on a set of observed covariates, assignment to the treatment is statistically independent of the untreated potential outcomes (for more details, see Imbens–Rubin 2015, Masten–Poirier 2018, Hollingworth–Wing 2020). In the synthetic control concept, we assume that the choice of which unit will be treated is random conditional on the choice of the donor pool, observable variables included as predictors, and unobservable variables captured by the path of the outcome variable during the pre-intervention period (Firpo–Possebom 2018).

The *stable unit treatment value assumption* (*SUTVA*) requires that the realised outcome for each unit depends only on the value of the treatment of that unit, and not on the treatment or outcome values of other units (for more details, see Abadie–Cattaneo 2018, Firpo–Possebom 2018, Imbens–Rubin 2015, Rubin 1980, Yu et al. 2020). This means that the units have no interactions and the impact of the intervention on the treated unit does not extend to the donor pool units.

In the context of the SCM, a key assumption is that the unobservable factors affecting the outcome variable and predictors remain unaffected by the intervention; therefore, their role remains unchanged in the post-treatment period (Abadie et al. 2010). There are no latent factors in the pre-treatment period that can materially affect the development of the outcome variable in the post-treatment period. Therefore, the donors and treated unit largely depend on the same factors that change over time both pre- and post-treatment. One assumes besides the intervention, no other major events during the study period have varying impacts on the treated and donor pool units (Yu et al. 2020). If some rare events did not affect a donor unit in the pre-treatment period, but appeared in the post-treatment period, they may distort the estimates (Hollingsworth–Wing 2020).

3. The multivariate synthetic control method estimator

The multivariate SCM estimator (MSCMT) is a generalised and expanded version of the standard SCM developed by Klößner–Pfeifer (2018). It is used here owing to the following features:

First, the MSCMT algorithm yields more reliable and accurate results than earlier standard SCM implementations because it checks various special cases at the beginning of the process; if any are present, faster and more accurate estimates are produced. Moreover, Becker–Klößner (2018) showed that the WNNLS algorithm used during MSCMT enables much more accurate and significantly faster calculations in the inner optimisation task than the functions widely used for this purpose in SCM analyses. By comparing 16 candidates, the recommendation is to perform the outer optimisation task using heuristic optimisation, as it performs much better than the average in finding the correct solution.

Another motivation for employing the MSCMT process is that it allows one to consider the entire time series of economic predictors. Thus, the information on the change in predictors over time can be used while producing the synthetic control.

Klößner–Pfeifer (2018) proved that if the last p observation of the outcome variable and economic predictors before treatment have a perfect fit, the estimated impact is unbiased. The better the fit in the pre-treatment period, the smaller the potential bias. If the pre-treatment period is long, the potential bias tends towards zero.

The MSCMT algorithm further enhanced by Becker–Klößner (2018) is used here. During the impact assessment, the following two optimisation tasks must be solved:

The inner optimisation task aims to ensure that the fit of the synthetic control on the economic predictors is as good as possible. Thus, the $v_1, \ldots, v_K > 0$ predictor weights are derived endogenously by minimising the following difference:

$$\Delta_X(v_1, \ldots, v_K, W) = \sqrt{\sum_{k=1}^{K} v_k \frac{1}{N_k} \sum_{n=1}^{N_k} \gamma_{k,n} \left(X_{k,n,1} - \sum_{j=2}^{J+1} X_{k,n,j} w_j \right)^2}$$

where $X_{k,n,j}$ is the value of the kth economic predictor in the nth pre-treatment period for the jth donor unit. The weight $\gamma_{k,n}$ allows the fitting of the kth economic predictor to count more in the periods nearer to the intervention than in earlier ones.

The outer optimisation task aims to ensure that the donor weights are chosen such that the synthetic control for the outcome variable is as close to the actual data as possible during the pre-treatment period. This can be achieved by minimising the following expression:

$$\Delta_Y(W) = \sqrt{\frac{1}{M^{pre}} \sum_{m=1}^{M^{pre}} \left(Y_{m,1} - \sum_{j=2}^{J+1} Y_{m,j} w_j \right)^2}$$

where M^{pre} pre-treatment period observations of the outcome variable are considered, and $Y_{m,j}$ denotes the outcome variable value in the case of the jth donor unit in the mth pre-treatment period.

Table A1

Covariates used for the estimation

Variable name	Variable content	Source	Time horizon
City size, location			
Stock of dwellings		[4]	2010
Population		[4]	2010
Shortest travel time to Budapest		[3]	2006, 2008, 2010
Housing market conditions			
Turnover rate of the stock of dwellings	Number of housing market transactions per year divided by the stock of dwellings	[4], [7]	2004Q1–2010Q4
Average square metre prices		[7]	2004Q1–2010Q4
Renewal of the stock of dwellings	Number of home completions in a year divided by the stock of dwellings.	[4]	2004Q1–2010Q4
Proportion of empty homes within the stock of dwellings		[4]	2011
Affordability of home purchases	Number of years it takes to purchase a 60-square metre home from net annual income based on the average square metre prices and incomes	[7], [4]	2004Q1–2010Q4
Level of economic development			
After-tax income per capita	Annual net income subject to personal income tax divided by the population of the settlement	[7], [4]	2004Q1–2010Q4
Equity per capita	Total equity of the companies registered in the settlement divided by the population.	[7], [4]	2004Q1–2010Q4
Export sales revenue per capita	Total annual export sales revenue of the companies registered in the settlement divided by the population	[7], [4]	2004Q1–2010Q4
Demographics			
Net migration per 1,000	Annual difference between incoming and outgoing migrants per 1,000 residents	[4]	2004Q1–2010Q4
Social aspects			
Number of people in employment within the population		[4]	2011
Proportion of community housing within the stock of dwellings		[4]	2010
Proportion of homes with all the modern conveniences within the stock of dwellings		[4]	2011
Number of hospital beds per thousand person	Number of hospital beds divided by 1,000 residents	[4]	2009

Note: Covariates in the table are used to estimate the models. Variables with italic text were excluded from the final model because of their low relevance based on the authors' first estimates.

Table A2

Summary statistics of the covariates for the pre- and post-treatment periods

Denomination	Pre-treatment period			
	minimum	average	maximum	st.dev.
Stock of dwellings, number	14,977	43,968	90,342	22,683
Population, person	33,720	100,630	208,016	50,555
Shortest travel time to Budapest, minute	42	110	174	42
Turnover rate of the stock of dwellings, %	0.8	1.1	1.2	0.1
Average square metre prices, HUF	137,445	174,489	213,585	23,724
Renewal of the stock of dwellings, %	0.4	0.9	1.6	0.3
Proportion of empty homes within the stock of dwellings, %	6	9	11	2
Affordability of home purchases, year	13	16	21	2
After-tax income per capita, HUF	583,082	653,201	771,767	56,293
Equity per capita, HUF	905	5,632	61,619	14,964
Export sales revenue per capita, HUF	366	1,790	10,441	2,450
Net migration per thousand person, number/thousand person	–4	1	6	3
Number of people in employment within the population, %	38	42	45	2
Proportion of community housing within the stock of dwellings, %	1	4	7	2
Proportion of homes with all the modern conveniences within the stock of dwellings, %	56	70	92	10
Number of hospital beds per thousand person, number/thousand person	8	16	29	5
Denomination	**Post-treatment period**			
	minimum	average	maximum	st.dev.
Stock of dwellings, number	15,201	45,289	95,578	23,826
Population, person	33,046	96,738	203,440	48,716
Shortest travel time to Budapest, minute	55	124	189	43
Turnover rate of the stock of dwellings, %	0.7	0.9	1.0	0.1
Average square metre prices, HUF	115,720	161,723	219,409	31,734
Renewal of the stock of dwellings, %	0.1	0.2	0.4	0.1
Proportion of empty homes within the stock of dwellings, %	6	9	11	2
Affordability of home purchases, year	8	11	15	2
After-tax income per capita, HUF	806,580	901,351	1,079,847	72,324
Equity per capita, HUF	1,470	5,256	34,510	8,163
Export sales revenue per capita, HUF	580	3,383	10,044	3,059
Net migration per thousand person, number/thousand person	–5	–1	4	2
Number of people in employment within the population, %	38	42	45	2
Proportion of community housing within the stock of dwellings, %	1	4	7	2
Proportion of homes with all the modern conveniences within the stock of dwellings, %	56	70	92	10
Number of hospital beds per thousand person, number/thousand person	8	17	29	5

Note: Summary statistics are calculated by averaging the quarterly values over the pre- and post-treatment periods for each element of the donor pool. For the pre-treatment period, quarterly values included in the estimation (last column of Table A1) are considered in the averaging process.

Table A3

Further results of the models' robustness analysis

Denomination	Baseline modell	Robustness analysis								
		Without economic predictors, with all the pre-treatment period values of the outcome variable	Shorter sample, pre-intervention period starts in 2005 Q1	Shorter sample, pre-intervention period starts in 2006 Q1	Excluding the donor with the largest weight, i.e. Szolnok	Excluding the donor with the second largest weight, i.e. Eger	Seed=50 for the random number generator	Genetic algorithm is used for outer-optimization	Testing anticipation effect: pre-intervention period ends in 2011 Q4	In-time placebo test: pre-intervention period ends in 2017 Q4
Donor weights										
Szolnok	28.23	23.89	43.69	0.55	0.00	9.09	28.23	29.10	33.36	28.64
Eger	16.27	13.93	24.29	18.62	0.00	0.00	16.27	15.50	0.00	7.75
Békéscsaba	13.82	13.34	9.35	20.06	20.98	20.63	13.82	15.17	15.28	24.90
Pécs	12.23	15.47	6.43	14.40	22.77	22.04	12.23	10.22	0.00	28.25
Tatabánya	10.09	11.59	7.37	12.31	14.15	12.64	10.09	10.04	0.00	2.98
Szekszárd	9.86	9.15	7.31	12.64	10.80	12.91	9.86	10.79	4.24	0.00
Veszprém	9.11	6.45	0.00	21.29	15.56	18.11	9.11	8.78	11.82	4.20
Szombathely	0.39	4.07	0.54	0.11	0.00	0.19	0.39	0.41	0.04	0.00
Debrecen	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.67	0.00
Kaposvár	0.00	0.00	0.00	0.00	6.72	3.83	0.00	0.00	0.00	0.00
Miskolc	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
Nyíregyháza	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	10.03	0.00
Szeged	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	21.58	0.00
Székesfehérvár	0.00	0.00	1.01	0.00	5.10	0.56	0.00	0.00	0.00	0.00
Zalaegerszeg	0.00	2.12	0.00	0.00	3.91	0.00	0.00	0.00	0.00	3.29

(Table continues on the next page.)

(Continued.)

Denomination	Base-line modell	Robustness analysis								
		Without economic predictors, with all the pre-treatment period values of the outcome variable	Shorter sample, pre-intervention period starts in 2005 Q1	Shorter sample, pre-intervention period starts in 2006 Q1	Excluding the donor with the largest weight, i.e. Szolnok	Excluding the donor with the second largest weight, i.e. Eger	Seed=50 for the random number generator	Genetic algorithm is used for outer-optimization	Testing anticipation effect: pre-intervention period ends in 2011 Q4	In-time placebo test: pre-intervention period ends in 2017 Q4
Variable weights										
Equity per capita	97.55	0.00	98.21	0.03	0.00	87.43	97.55	98.17	0.00	0.00
House price index	2.41	100.00	1.78	0.00	85.84	12.02	2.41	1.81	0.03	93.62
Proportion of community housing within the stock of dwellings	0.02	0.00	0.00	99.97	0.52	0.17	0.02	0.01	0.01	0.10
Average square metre prices	0.01	0.00	0.00	0.00	0.00	0.04	0.01	0.00	99.95	
Shortest travel time to Budapest	0.01	0.00	0.00	0.00	0.26	0.04	0.01	0.00	0.00	0.02
Number of hospital beds per thousand person	0.00	0.00	0.00	0.00	0.16	0.00	0.00	0.00	0.00	0.00
After-tax income per capita	0.00	0.00	0.00	0.00	0.00	0.30	0.00	0.00	0.00	4.82
Export sales revenue per capita	0.00	0.00	0.00	0.00	12.51	0.00	0.00	0.00	0.00	0.00
Stock of dwellings	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Renewal of the stock of dwellings	0.00	0.00	0.00	0.00	0.71	0.00	0.00	0.00	0.00	0.76
Turnover rate of the stock of dwellings	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.68
Population	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Proportion of empty homes within the stock of dwellings	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Affordability of home purchases	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Note: The results in the table refer to the model with the smallest RMSPE value during the pre-intervention period.

REFERENCES

ABADIE, A. (2021): Using synthetic controls: Feasibility, data requirements, and methodological aspects *Journal of Economic Literature* 59 (2): 391–425. https://doi.org/10.1257/jel.20191450

ABADIE, A.–GARDEAZABAL, J. (2003): The economic costs of conflict: A case study of the Basque Country *American Economic Review* 93 (1): 112–132. https://doi.org/10.1257/000282803321455188

ABADIE, A.–CATTANEO, M. (2018): Econometric methods for program evaluation *Annual Review of Economics* 10 (1): 465–503. https://doi.org/10.1146/annurev-economics-080217-053402

ABADIE, A.–DIAMOND, A.–HAINMUELLER, J. (2010): Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program *Journal of the American Statistical Association* 105 (490): 493–505. https://doi.org/10.1198/jasa.2009.ap08746

ABADIE, A.–DIAMOND, A.–HAINMUELLER, J. (2015): Comparative politics and the synthetic control method *American Journal of Political Science* 59 (2): 495–510. https://doi.org/10.1111/ajps.12116

ABADIE, A.–L'HOUR, J. (2021): A penalized synthetic control estimator for disaggregated data *Journal of the American Statistical Association* 116 (536): 1817–1834. https://doi.org/10.1080/01621459.2021.1971535

ABRAHAM, J. M.–HENDERSHOTT, P. H. (1996): Bubbles in Metropolitan housing markets *Journal of Housing Research* 7 (2): 191–207. https://doi.org/10.3386/w4774

ADAMS, Z.–FÜSS, R. (2010): Macroeconomic determinants of international housing markets *Journal of Housing Economics* 19 (1): 38–50. https://doi.org/10.1016/j.jhe.2009.10.005

AGNELLO, L.–SCHUKNECHT, L. (2011): Booms and busts in housing markets: Determinants and implications *Journal of Housing Economics* 20 (3): 171–190. https://doi.org/10.1016/j.jhe.2011.04.001

AGNEW, K.–LYONS, R. C. (2018): The impact of employment on housing prices: Detailed evidence from FDI in Ireland *Regional Science and Urban Economics* 70: 174–189. https://doi.org/10.1016/j.regsciurbeco.2018.01.011

ANTALÓCZY, K.–BIRIZDÓ, I.–SASS, M. (2022): Local investment promotion in a Hungarian medium-sized town and the implications of the COVID pandemic *Regional Statistics* 12 (1): 27–50. https://doi.org/10.15196/RS120104

APERGIS, N.–REZITIS, A. (2003): Housing prices and macroeconomic factors in Greece: Prospects within the EMU *Applied Economics Letters* 10 (9): 561–565. https://doi.org/10.1080/1350485032000100260

BACCINI, L.–LI, Q.–MIRKINA, I. (2014): Corporate tax cuts and foreign direct investment *Journal of Policy Analysis and Management* 33 (4): 977–1006. https://doi.org/10.1002/pam.21786

BAFFOE-BONNIE, J. (1998): The dynamic impact of macroeconomic aggregates on housing prices and stock of houses: A national and regional analysis *The Journal of Real Estate Finance and Economics* 17: 179–197. https://doi.org/10.1023/A:1007753421236

BANAI, Á.–VÁGÓ, N.–WINKLER, S. (2017): The MNB's house price index methodology *MNB Occasional Papers* 2017/127. Central Bank of Hungary.

BANAI, Á.–VÁGÓ, N.–WINKLER, S. (2018): Measuring heterogeneity of house price developments in Hungary, 1990–2016 *Acta Oeconomica* 68 (3): 377–414. https://doi.org/10.1556/032.2018.68.3.4

BARONE, G.–MOCETTI, S. (2014): Natural disasters, growth and institutions: A tale of two earthquakes *Journal of Urban Economics* 84 (C): 52–66. https://doi.org/10.1016/j.jue.2014.09.002

BECKER, M.–KLÖßNER, S. (2018): Fast and reliable computation of generalized synthetic controls *Econometrics and Statistics* 5 (C): 1–19. https://doi.org/10.1016/j.ecosta.2017.08.002

BECKER, M.–KLÖßNER, S.–PFEIFER, G. (2018): Cross-validating synthetic controls *Economics Bulletin* 38 (1): 603–609.

BECKER, M.–PFEIFER, G.–SCHWEIKERT, K.–PFEIFER, G. G. (2021): Price effects of the Austrian fuel price fixing act: A synthetic control study CESifo Working Paper Series 8819, CESifo. https://ideas.repec.org/p/ces/ceswps/_8819.html

BELKE, A.–KEIL, J. (2018): Fundamental determinants of real estate prices: A panel study of German regions *International Advances in Economic Research* 24 (1): 25–45. https://doi.org/10.1007/s11294-018-9671-2

BEN-MICHAEL, E.–FELLER, A.–ROTHSTEIN, J. (2021): The augmented synthetic control method *Journal of the American Statistical Association* 116 (536): 1789–1803. https://doi.org/10.1080/01621459.2021.1929245

BÉRES, A.–JABLONSZKY, GY.–LAPOSA, T.–NYIKOS, GY. (2019): Spatial econometrics: transport infrastructure development and real estate values in Budapest *Regional Statistics* 9 (2): 89–104. https://doi.org/10.15196/RS090202

BILLMEIER, A.–NANNICINI, T. (2013): Assessing economic liberalization episodes: A synthetic control approach *The Review of Economics and Statistics* 95 (3): 983–1001. https://doi.org/10.1162/REST_a_00324

BOHN, S.–LOFSTROM, M.–RAPHAEL, S. (2014): Did the 2007 legal Arizona workers act reduce the State's unauthorized immigrant population? *The Review of Economics and Statistics* 96 (2): 258–269. https://doi.org/10.1162/REST_a_00429

BOTOSARU, I.–FERMAN, B. (2019): On the role of covariates in the synthetic control method *Econometrics Journal* 22 (2): 117–130. https://doi.org/10.1093/ectj/utz001

BOXALL, P. C.–WING, H. C.–MCMILLAN, M. L. (2005): The impact of oil and natural gas facilities on rural residential property values: A spatial hedonic analysis *Resource and Energy Economics* 27 (3): 248–269. https://doi.org/10.1016/j.reseneeco.2004.11.003

BRISSIMIS, S. N.–VLASSOPOULOS, T. (2009): The interaction between mortgage financing and housing prices in Greece *Journal of Real Estate Economics* 39: 146–164. https://doi.org/10.1007/s11146-008-9109-3

BRONZINI, R.–MOCETTI, S.–MONGARDINI, M. (2020): The economic effects of big events: Evidence from the Great Jubilee 2000 in Rome *Journal of Regional Science* 60 (4): 801–822. https://doi.org/10.1111/jors.12485

CAMPOS, N. F.–CORICELLI, F.–MORETTI, L. (2019): Institutional integration and economic growth in Europe *Journal of Monetary Economics* 103: 88–104.

https://doi.org/10.1016/j.jmoneco.2018.08.001

CAPOZZA, D.–HENDERSHOTT, P.–MACK, C.–MAYER, C. (2002): Determinants of real house price dynamics *NBER Working Papers* No. 9262. National Bureau of Economic Research, Inc, Cambridge, MA. https://doi.org/10.3386/w9262

CAVALLERI, M. C.–COURNÈDE, B.–ÖZSÖĞÜT, E. (2019): How responsive are housing markets in the OECD? National level estimates *OECD Economics Department Working Papers* No. 1589. OECD Publishing, Paris. https://doi.org/10.1787/18151973

CHERNOZHUKOV, V.–WÜTHRICH, K.–ZHU, Y. (2021): An exact and robust conformal inference method for counterfactual and synthetic controls *Journal of the American Statistical Association* 116 (536): 1849–1864. https://doi.org/10.1080/01621459.2021.1920957

COFFMAN, M.–NOY, I. (2012): Hurricane Iniki: Measuring the long-term economic impact of a natural disaster using synthetic control *Cambridge University Press* 17 (2): 187–205. https://doi.org/10.1017/S1355770X11000350

CURRIE, J.–DAVIS, L.–GREENSTONE, M.–WALKER, R. (2015): Environmental health risks and housing values: Evidence from 1,600 toxic plant openings and closings *American Economic Review* 105 (2): 678–709. https://doi.org/10.1257/aer.20121656

DAI, X.–BAI, X.–XU, M. (2016): The influence of Beijing rail transfer stations on surrounding housing prices *Habitat International* 55: 79–88. https://doi.org/10.1016/j.habitatint.2016.02.008

DAY, C. (2018): Australia's growth in households and house prices *Australian Economic Review* 51 (4): 502–511. https://doi.org/10.1111/1467-8462.12280

DE SOUSA, C. (2005): Policy performance and Brownfield redevelopment in Milwaukee, Wisconsin *The Professional Geographer* 57 (2): 312–327. https://doi.org/10.1111/j.0033-0124.2005.00480.x

DE SOUSA, C. A.–WU, C.–WESTPHAL, L. M. (2009): Assessing the effect of publicly assisted Brownfield redevelopment on surrounding property values *Economic Development Quarterly* 23 (2): 95–110. https://doi.org/10.1177/0891242408328379

DE VOR, F.–DE GROOT, H. L. F. (2011): The impact of industrial sites on residential property values: A hedonic pricing analysis from the Netherlands *Regional Studies* 45 (5): 609–623. http://dx.doi.org/10.1080/00343401003601925

DONOHUE, J. J.–ANEJA, A.–WEBER, K. D. (2019): Right-to-carry laws and violent crime: A comprehensive assessment using panel data and a state-level synthetic control analysis *Journal of Empirical Legal Studies* 16 (2): 198–247. https://doi.org/10.1111/jels.12219

DOUDCHENKO, N.–IMBENS, G. W. (2016): Balancing, regression, difference-in-differences and synthetic control methods: A synthesis *NBER Working Papers* No. 22791., National Bureau of Economic Research, Inc., Cambridge, MA. https://doi.org/10.3386/w22791

DUBE, A.–ZIPPERER, B. (2015): Pooling multiple case studies using synthetic controls: An application to minimum wage policies *IZA Discussion Papers* No. 8944., Institute for the Study of Labor (IZA), Bonn, Germany.

DUCA, J. V.–MUELLBAUER, J.–MURPHY, A. (2021): What drives house price cycles? International experience and policy issues *Journal of Economic Literature* 59 (3): 773–864. https://doi.org/10.1257/jel.20201325

DUSTMANN, C.–SCHONBERG, U.–STUHLER, J. (2017): Labor supply shocks, native wages, and the adjustment of local employment *The Quarterly Journal of Economics* 132 (1): 435–483. http://dx.doi.org/10.1093/qje/qjw032

EREN, O.–OZBELIK, S. (2017): What do right-to-work laws do? Evidence from a synthetic control method analysis *Journal of Policy Analysis and Management* 35 (1): 173–194. https://doi.org/10.1002/pam.21861

FALKENHALL, B.–MANSSON, J.–TANO, S. (2020): Impact of VAT reform on Swedish restaurants: A synthetic control group approach *Scandinavian Journal of Economics* 122 (2): 824–850. https://doi.org/10.1111/sjoe.12340

FERMAN, B.–PINTO, C.–POSSEBOM, V. (2020): Cherry picking with synthetic controls *Journal of Policy Analysis and Management* 39 (2): 510–532. https://doi.org/10.1002/pam.22206

FERRACCI, M.–JOLIVET, G.–VAN DEN BERG, G. (2014): Evidence of treatment spillovers within markets *The Review of Economics and Statistics* 96 (5): 812–823. https://doi.org/10.1162/REST_a_00428

FIRPO, S.–POSSEBOM, V. (2018): Synthetic control method: Inference, sensitivity analysis and confidence sets *Journal of Causal Inference* 6 (2): 1–26. https://doi.org/10.1515/jci-2016-0026

FITZPATRICK, T.–MCQUINN, K. (2004): House prices and mortgage credit: empirical evidence for Ireland *Central Bank and Financial Services Authority of Ireland Research Technical Paper* 5/RT/04. Dublin, Ireland.

GABRIEL, S.–IACOVIELLO, M.–LUTZ, C. (2021): A crisis of missed opportunities? Foreclosure costs and mortgage modification during the great recession *Review of Financial Studies* 34 (2): 864–906. https://doi.org/10.1093/rfs/hhaa059

GAUTIER, P. A.–SIEGMANN, A.–VAN VUUREN, A. (2009): Terrorism and attitudes towards minorities: The effect of the Theo van Gogh murder on house prices in Amsterdam *Journal of Urban Economics* 65 (2): 113–126. https://doi.org/10.1016/j.jue.2008.10.004

GERLACH, S.–PENG, W. (2005): Bank lending and property prices in Hong Kong *Journal of Banking & Finance* 29 (2): 461–481. https://doi.org/10.1016/j.jbankfin.2004.05.015

GIMENO, R.–MARTÍNEZ-CARRASCAL, C. (2006): The interaction between house prices and loans for house purchase. The Spanish case *Banco de España Documentos de Trabajo* No. 0605. BANCO DE ESPAÑA, Madrid.

GLAESER, E. L.–GYOURKO, J.–SAKS, R. E. (2005): Why have housing prices gone up? *American Economic Review* 95 (2): 329–333.

GLAESER, E. L.–GYOURKO, J.–SAIZ, A. (2008): Housing supply and housing bubbles *Journal of Urban Economics* 64 (2): 198–217. https://doi.org/10.1016/j.jue.2008.07.007

GONG, Y.–YAO, Y. (2022): Demographic changes and the housing market *Regional Science and Urban Economics* 95: 103734. https://doi.org/10.1016/j.regsciurbeco.2021.103734

GRABOVAC, I.–HOCHFELLNER, L.–RIEGER, M.–JEWEL, J.–SNELL, A.–WEBBER, A.–STÜGER, H. P.–SCHINDLER, K.–MIKKELSEN, B.–DORNER, T. (2018): Impact of Austria's 2009 trans fatty acids regulation on all-cause, cardiovascular and coronary heart disease mortality *The European Journal of Public Health* 28 (2): 4–9. http://dx.doi.org/10.1093/eurpub/cky147

GRISLAIN-LETRÉMY, C.–KATOSSKY, A. (2014): The impact of hazardous industrial facilities on housing prices: A comparison of parametric and semiparametric hedonic price models *Regional Science and Urban Economics* 49: 93–107. https://doi.org/10.1016/j.regsciurbeco.2014.09.002

HANINGER, K.–MA, L.–TIMMINS, C. (2017): The value of Brownfield remediation *Journal of the Association of Environmental and Resource Economists* 4 (1): 197–241. https://doi.org/10.1086/689743

HILBER, C. A. L.–VERMEULEN, W. (2016): The impact of supply constraints on house prices in England *Economic Journal* 126 (591): 358–405. https://doi.org/10.1111/ecoj.12213

HILLER, N.–LERBS, O. W. (2016): Aging and urban house prices *Regional Science and Urban Economics* 60: 276–291. https://doi.org/10.1016/j.regsciurbeco.2016.07.010

HIMMELBERG, C.–MAYER, C.–SINAI, T. (2005): Assessing high house prices: Bubbles, fundamentals and misperceptions *Journal of Economic Perspectives* 19 (4): 67–92. https://doi.org/10.1257/089533005775196769

HINRICHS, P. (2012): The effects of affirmative action bans on college enrollment, educational attainment, and the demographic composition of universities *Review of Economics and Statistics* 94 (3): 712–722. https://doi.org/10.1162/REST_a_00170

HOLLINGSWORTH, A.–WING, C. (2020): Tactics for design and inference in synthetic control studies: An applied example using high-dimensional data *SSRN*. http://dx.doi.org/10.2139/ssrn.3592088

IMBENS, G.–RUBIN, D. (2015): *Causal inference for statistics, social, and biomedical sciences: An introduction* Cambridge University Press, Cambridge.

KAUL, A.–KLÖßNER, S.–PFEIFER, G.–SCHIELER, M. (2021): Standard synthetic control methods: The case of using all preintervention outcomes together with covariates *Journal of Business & Economic Statistics* 40 (3): 1362–1376. https://doi.org/10.1080/07350015.2021.1930012

KELLOGG, M.–MOGSTAD, M.–POULIOT, G.–TORGOVITSKY, A. (2021): Combining matching and synthetic control to tradeoff biases from extrapolation and interpolation *Journal of the American Statistical Association* 116 (536): 1804–1816. https://doi.org/10.1080/01621459.2021.1979562

KIM, J.–LEE, S. (2022): Foreign direct investment and housing prices: Evidence from South Korea *International Economic Journal* 36 (2): 247–262. https://doi.org/10.1080/10168737.2022.2055108

KLEVEN, H. J.–LANDAIS, C.–SAEZ, E. (2013): Taxation and international migration of superstars: Evidence from European football market *American Economic Review* 103 (5): 1892–1924. https://doi.org/10.1257/aer.103.5.1892

LEE, J.–JUNG, H. (2020): Demographic shifts, macroprudential policies, and house prices *BIS Working Papers* No. 914., Bank for International Settlements.

LINDO, J. M.–PACKHAM, A. (2017): How much can expanding access to long-acting reversible contraceptives reduce teen birth rates? *American Economic Journal: Economic Policy* 9 (3): 348–376. https://doi.org/10.1257/pol.20160039

LINN, J. (2013): The effect of voluntary brownfields programs on nearby property values: Evidence from Illinois *Journal of Urban Economics* 78: 1–18. https://doi.org/10.1016/j.jue.2013.04.002

LIU, S. (2015): Spillovers from universities: Evidence from the land-grant program *Journal of Urban Economics* 87 (C): 25–41. https://doi.org/10.1016/j.jue.2015.03.001

MARTINELLI, C.–VEGA, M. (2019): The economic legacy of general Velasco: Long-term consequences of interventionism *Economía* 42 (84): 102–133. https://doi.org/10.18800/economia.201902.005

MASTEN, M. A.–POIRIER, A. (2018): Identification of treatment effects under conditional partial independence *Econometrica* 86 (1): 317–351. https://doi.org/10.3982/ECTA14481

MIKELBANK, B.A. (2004): Spatial analysis of the relationship between housing values and investments in transportation infrastructure *The Annals of Regional Science* 38: 705–726. https://doi.org/10.1007/s00168-003-0187-0

MOHAMMAD, S.–GRAHAM, D.–MELO, P. C.–ANDERSON, R. (2013): A meta-analysis of the impact of rail projects on land and property values *Transportation Research Part A: Policy and Practice* 50: 158–170. https://doi.org/10.1016/j.tra.2013.01.013

MONASTIRIOTIS, V.–ZILIC, I. (2020): The economic effects of political disintegration: Lessons from Serbia and Montenegro *European Journal of Political Economy* 65: 101938. https://doi.org/10.1016/j.ejpoleco.2020.101938

OPATRNY, M. (2021): The impact of the Brexit vote on UK financial markets: A synthetic control method approach *Empirica* 48 (2): 559–587. https://doi.org/10.1007/s10663-020-09481-7

PERI, G.–YASENOV, V. (2019): The labor market effects of a refugee wave: Synthetic control method meets the Mariel boatlift *Journal of Human Resources* 54 (2): 267–309. https://doi.org/10.3368/jhr.54.2.0217.8561R1

QUIGLEY, J. M.–RAPHAEL, S. (2005): Regulation and the high cost of housing in California *American Economic Review* 95 (2): 323–328. https://doi.org/10.1257/000282805774670293

ROSENBAUM, P. R.–RUBIN, D. B. (1983): The central role of the propensity score in observational studies for causal effects *Biometrika* 70 (1): 41–55. https://doi.org/10.1093/biomet/70.1.41

ROSENBAUM, P. R.–RUBIN, D. B. (1985): Constructing a control group using multivariate matched sampling methods that incorporate the propensity score *The American Statistician* 39 (1): 33–38. https://doi.org/10.1080/00031305.1985.10479383

ROUWENDAL, J.–PETRAT, A. (2022): Mortgage underwriting and house prices: Evidence from the 2011 change in the Dutch Code of Conduct for mortgage loans *Real Estate Economics* 50 (4): 1141–1159. https://doi.org/10.1111/1540-6229.12372

RUBIN, D. B. (1980): Randomization analysis of experimental data in the Fisher randomization test *Journal of the American Statistical Association* 75 (371): 591–593. https://doi.org/10.1080/01621459.1980.10477512

SAIZ, A. (2010): The geographic determinants of housing supply *The Quarterly Journal of Economics* 125 (3): 1253–1296. https://doi.org/10.1162/qjec.2010.125.3.1253

SAKS, R. E. (2008): Job creation and housing construction: Constraints on metropolitan area employment growth *Journal of Urban Economics* 64 (1): 178–195. https://doi.org/10.1016/j.jue.2007.12.003

STEARNS, J. (2015): The effects of paid maternity leave: Evidence from temporary disability insurance *Journal of Health Economics* 43 (490): 85–102. https://doi.org/10.1016/j.jhealeco.2015.04.005

TAKÁTS, E. (2012): Aging and house prices *Journal of Housing Economics* 21 (2): 131–141. https://doi.org/10.1016/j.jhe.2012.04.001

TAN, R.–HE, Q.–ZHOU, K.–XIE, P. (2019): The effect of new metro stations on local land use and housing prices: The case of Wuhan, China *Journal of Transport Geography* 79: 102488. https://doi.org/10.1016/j.jtrangeo.2019.102488

TAYLOR, J. B. (2009): The financial crisis and the policy responses: An empirical analysis of what went wrong *NBER Working Paper* No.14631. National Bureau of Economic Research, Cambridge, MA. https://doi.org/10.3386/w14631

WANG, Y.–KINUGASA, T. (2022): The relationship between demographic change and house price: Chinese evidence *International Journal of Economic Policy Studies* 16 (1): 43–65. https://doi.org/10.1007/s42495-021-00068-z

WOO, A.–LEE, S. (2016): Illuminating the impacts of brownfeld redevelopments on neighboring housing prices: Case of Cuyahoga County, Ohio in the US. *Environment and Planning A: Economy and Space* 48 (6): 1107–1132. https://doi.org/10.1177/0308518X16636380

XU, L.–TANG, B. (2014): On the determinants of UK house prices *International Journal of Economics and Research* 5 (2): 57–64.

XU, Y. (2017): Generalized synthetic control method: Causal inference with interactive fixed effects Models *Political Analysis* 25 (1): 57–76. https://doi.org/10.1017/pan.2016.2

YU, V.–WYATT, S.–WOODALL, M.–MAHMUD, S.–KLAIRE, V.–BAILEY, K.–MOHAMMED, M. A. (2020): Hospital admissions after vertical integration of general practices with an acute hospital: A retrospective synthetic matched controlled database study *British Journal of General Practice* 70 (699): 705–713. https://doi.org/10.3399/bjgp20x712613

INTERNET SOURCES

KLÖßNER, S.–PFEIFER, G. (2018): Synthesizing cash for clunkers: Stabilizing the car market, hurting the environment? MPRA Paper No. 88175. https://mpra.ub.uni-muenchen.de/88175 (downloaded: August 2022)

PAVLÍNEK, P.–ALÁEZ-ALLER, R.–GIL-CANALETA, C.–ULLIBARRI-ARCE, M. (2017): Foreign direct investment and the development of the automotive industry in Eastern and Southern Europe *European Trade Union Institute, Working Paper* 2017.03. https://www.etui.org/sites/default/files/17%20WP%202017%2003%20Foreign%20Direct%20Investment%20and%20the%20development%20EN%20Web%20version.pdf (downloaded: August 2022)

WEBSITES/DATABASES

[1] AZO MATERIALS: https://www.azom.com/article.aspx?ArticleID=22236 (downloaded: August 2020)

[2] EUROSTAT: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Manufacturing_statistics_-_NACE_Rev._2#Sectoral_analysis (downloaded: August 2020)

[3] GEOX KFT.: (downloaded: August 2020)

[4] HUNGARIAN CENTRAL STATISTICAL OFFICE (HCSO): (downloaded: August 2020)

[5] MCKINSEY&COMPANY: https://www.mckinsey.com/capabilities/sustainability/our-insights/spotting-green-business-opportunities-in-a-surging-net-zero-world/transition-to-net-zero/road-mobility (downloaded: August 2020)

[6] NATIONAL REGIONAL DEVELOPMENT AND SPATIAL PLANNING INFORMATION SYSTEM (in Hungarian: TeIR): (downloaded: August 2020)

[7] NATIONAL TAX AND CUSTOMS ADMINISTRATION (in Hungarian: NAV): (downloaded: August 2020)