buildings

Article

Forecasting Total and Type-Specific Non-Residential Building Construction Spending: The Case Study of the United States and Lessons Learned

Xingrui Zhang [1,*], Yunpeng Wang [2], Shuai Xu [3], Eunhwa Yang [4] and Lingxiao Meng [1]

[1] China Construction Infrastructure Corp., Beijing 100029, China; muluo690119@163.com
[2] 2nd Engineering Bureau, China State Construction Engineering Corp., Beijing 100071, China; wangyunpeng871@cscec.com
[3] Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, CA 90033, USA; sxu75374@usc.edu
[4] School of Building Construction, Georgia Institute of Technology, Atlanta, GA 30332, USA; eunhwa.yang@design.gatech.edu
* Correspondence: xzhng147@gatech.edu or xingrui2015@hotmail.com

Abstract: Forecasting construction spending is important for civil engineering practitioners to make business decisions. Currently, the main body of forecasting literature pertains exclusively to aggregate construction investment, such as total construction spending (TTLCON), private construction spending, or residential construction spending. But type-specific construction spending, such as that for education, healthcare, and religion, had yet to be explored using forecasting techniques. This case study presents a viable procedure by which aggregate and type-specific non-residential construction can be forecasted. The procedure that involves the use of the Granger causality test and the Vector Autoregression (VAR) model proved to be able to provide an accurate forecast pre-COVID-19, with some accuracy even during the COVID-19 pandemic period. Lessons learned include the following: (1) effort should be diverted towards model interpretation, as the impulse–response trial yields results conforming to current well-established empirical evidence; (2) a type-specific approach should be adopted when analyzing construction spending, as different types of construction spending react differently to potential indicators; and (3) complex models incorporating multiple indicators should be used to generate a forecast, as a complex model has a higher chance of containing parameters explanatory of the target variable's features during the testing period.

Keywords: construction spending; time-series analysis; forecasting

Citation: Zhang, X.; Wang, Y.; Xu, S.; Yang, E.; Meng, L. Forecasting Total and Type-Specific Non-Residential Building Construction Spending: The Case Study of the United States and Lessons Learned. *Buildings* **2024**, *14*, 1317. https://doi.org/10.3390/buildings14051317

Academic Editor: Jorge Lopes

Received: 3 April 2024
Revised: 25 April 2024
Accepted: 30 April 2024
Published: 7 May 2024

1. Introduction

Construction spending, which reflects the value of construction put in place, has long been recognized as an important economic indicator. Forecasting construction spending is essential for aiding the decision-making of both civil engineering practitioners and policymakers. Notably, construction spending comoves with other economic indicators, including Gross Domestic Product and consumption [1]. Currently, the body of forecasting studies on construction pertains exclusively to aggregate construction spending, such as total residential/non-residential construction, as the two broad categories tend to impact employment conditions most. However, type-specific construction spending, such as that for education and religion, has yet to be explored. It can be reasonably inferred from the body of knowledge that different kinds type-specific construction spending display different correlations with indicators and should not be generalized.

In addition to that, the main body of forecasting literature in the field of construction economics focuses mostly on statistically significant correlations generated via black-box processes, and puts little effort into model interpretation. For example, Abiniangerabi et al., 2017 [2], Ashuri et al., 2012 [3], and Ahmadi and Shahandashti 2017 [4] use significant

F-statistics via the Granger procedure to assert correlations between construction spending and the Architectural Billing Index and economic indicators, but do not show further justification including whether the correlation is positive or negative. Such an approach obscures the qualitative implications that should have been drawn.

Therefore, the construction forecasting literature requires an update, taking into consideration (1) the forecastability of type-specific construction spending, and (2) model interpretation, which should show how various kinds of construction spending respond differently to economic conditions. This study starts with a summary of past work, then measures the Granger correlation between type-specific construction spending and economic indicators extracted from Federal Reserve Economic Data (FRED) and the Engineering News Record (ENR). A multivariate time-series model will then be implemented, and an extensive attempt at model interpretation via impulse–response functions will be made prior to generating a reasonable forecast.

2. Literature Review

2.1. Construction Spending Affects Employment

From an employment perspective, construction has traditionally been used to stimulate employment/job creation [5,6], albeit with varying levels of success. Ball and Wood 1995 [7] surveyed the employment generation effect of construction expenditure in the United Kingdom but found only a weak correlation, though such a contradiction with theoretical predictions is possibly due to poor data quality. On the other hand, Hassan 2017 [8] used multiple time-series simulation models to observe both local dynamics and long-term relationships between construction spending and employment, demonstrating strong causal relationships between the two parameters. Simonson's 2013 [9] presentation indicated a generally positive linear correlation between construction spending and construction employment between 2006 and 2016, with the lowest point of construction spending in early 2011 and that of construction employment somewhere between 2010 and 2011. This time-lagged relationship was further substantiated by Zhang et al., 2023 [10], whose study showed a 95% positive response in residential construction spending to prior increases in employment, with a lag of 10–11 months.

Construction spending also stimulates employment in neighboring industries, such as architecture, and causal relationships have been observed between construction spending and the Architecture Billings Index (ABI) [2,11,12]. Byun 2010 [13] noted that construction spending relates to "seemingly unrelated industries" as well as construction employment, suggesting a correlation between construction spending and the overall employment condition in a region. Generally speaking, in developed countries like the United States [14] and Singapore [15], construction spending is known to be strongly and positively correlated with productivity growth.

2.2. Construction Spending Affects Economic Strength

Construction spending is also known to be correlated with economic strength. Tse and Ganesan 1997 [16] used the Granger causality test on Hong Kong's time-series data and concluded that Gross Domestic Product (GDP) in Hong Kong leads construction spending. Davis and Heathcode 2005 focused on residential construction and housing value and concluded that the housing value and supply comoves with GDP and consumption [1]. Ahmadi and Shahandashti 2017 [4] utilized a combination of temporal and long-term causality tests and found that construction spending is a leading indicator of GDP in 18 of the States in the U.S., with 9 States seeing cointegration relationships that do not attenuate over time.

2.3. Gaps in Knowledge to Be Addressed

To date, most research on construction spending focuses on aggregate measures, such as total construction value added [4], total non-residential construction spending and institutional construction spending [2], and residential construction spending (PRRESCON) [10].

These time series are analyzed alongside unemployment and GDP as national economic indicators. However, data of type-specific construction spending, such as healthcare, education, religion, office, and lodging have not been extensively explored in a time-series fashion.

It can be reasonably inferred from the current body of empirical evidence that different types of construction spending respond differently to the same indicators. The most representative example is the relationship between construction spending and construction cost. The value of construction for a particular project is directly related to the construction cost, particularly in residential development [17]. However, the aggregate construction value of all construction carried out in an area may not correlate as directly with construction cost, as different types of development react differently to the fluctuation of construction cost. In housing, high construction costs typically lead to higher housing values [18,19]. For residential/housing development, changes in construction cost may precisely be in response to demand fluctuations, as construction cost is known to be strongly correlated with macroeconomic conditions that also impact housing prices [3,20,21]; the causal relationship between housing value and construction spending was found to be bidirectional [22], and positive correlation was discovered between construction cost, residential construction spending, and housing value [10]. On the other hand, there are types of development that are negatively impacted by construction cost escalations. Sectors like educational construction, which largely reflect non-profit endeavors, may be constrained by resource limitations [23]. Various types of construction spending may also respond differently to the disruption of COVID-19, as stated in Alsharef et al., 2021; in interviews with personnels of various sectors, the disruption of COVID-19 involves a mix of adverse effect, risk management, and opportunities [24]. While the disruption of COVID-19 will negatively impact some aspects of the construction industry by stressing workers' health and project finance [25,26], it might also be conducive to opportunities in some other sectors, such as the industry of care [27].

On top of that, the main body of forecasting literature pertaining to construction economics rely on a univariate time-series model [28], the Granger causality test and cointegration test [2–4,29], or deep learning techniques [30,31], with result verification performed by generating a forecast and then comparing the forecast with testing data.

All of the aforementioned techniques (Granger, cointegration, and neural network) employed should be considered as "black-box processes", since no interpretable results can be generated from those tests save for statistically significant F/T-statistics and a forecast to be compared with the testing data. However, correlation does not equate causality, and all of the aforementioned studies fail to answer the questions of (1) whether the correlation is positive or negative, (2) to what extent the correlation time is lagged, and (3) whether the regression features are qualitatively significant and yield practical implication, or correlation is simply in statistical form.

If the aforementioned questions remain unanswered, qualitative contributions, such as policy implications, cannot be made. Therefore, the construction economics forecasting literature is in need of an update that puts emphasis on model interpretation, policy implications, and type-specific construction.

Therefore, this case study presents an analysis of both aggregate and type-specific non-residential construction spending. The empirical results generated can be used to support other related studies while, at the same time, demonstrating the importance of model interpretation and taking a type-specific approach.

3. Material and Methods

3.1. Data Used

The data utilized in this case study were extracted from the Engineering News Record (Building Cost Index/BCI and Construction Cost Index/CCI) and the Federal Reserve Economic Data (FRED). The abbreviations adopted for specific datasets are aligned with FRED's nomenclature. These abbreviations, along with corresponding indicators, are summarized in Table 1. It is important to note that the data concerning residential construction spending (PRRESCON) are not included in the analysis, as this paper focuses exclusively on non-residential construction spending. However, PRRESCON is mentioned here as it will be referenced in subsequent sections for reasoning and explanatory purposes. All datasets employed in this paper were obtained as monthly raw values. The training period for model construction was set from January 1993 (M1 1993) to December 2017 (M12 2017), and the testing period was set from January 2018 (M1 2018) to December 2022 (M12 2022). Units of measurement are also provided in Table 1. More importantly, whether annual seasonality is present within the time series is also indicated. In order to maintain consistency, all data used are without seasonality adjustment. Even in their raw form, not all datasets have a consistent annual seasonality pattern.

Table 1. Summary of FRED and ENR data used.

Full Name	Abbreviation	Unit of Measurement	Annual Seasonality
Potential Leading Indicators			
All Building Construction Employees	CEU2023600001	Thousands of Persons	Yes
All Employees, Heavy Civil	CEU2023700001	Thousands of Persons	Yes
Average Construction Wage	AHECONS	Dollars Per Hour	No
New Housing Starts	HOUSTNSA	Thousands of Units	Yes
Unemployment	UNRATENSA	Percent	Yes
All Employees, Construction	CEU2000000001	Thousands of Persons	Yes
Construction Cost Index (ENR)	CCI	Standardized Units	No
Building Cost Index (ENR)	BCI	Standardized Units	No
Target Variables			
Private Office Construction Spending	PROFCON	Millions of Dollars	Yes
Private Religious Construction Spending	PRRELCON	Millions of Dollars	Yes
Total Construction Spending	TTLCON	Millions of Dollars	Yes
Private Non-Residential Construction Spending	PNRESCON	Millions of Dollars	Yes
Private Lodging Construction Spending	PLODGCON	Millions of Dollars	Yes
Private Commercial Construction Spending	PRCOMCON	Millions of Dollars	Yes
Private Educational Construction Spending	PREDUCON	Millions of Dollars	Yes
Private Healthcare Construction Spending	PRHLTHCON	Millions of Dollars	Yes
Private Manufacture Construction Spending	PRMFGCON	Millions of Dollars	Yes
Private Residential Construction Spending	PRRESCON	Millions of Dollars	Yes

In the subsequent sections of this paper, the term "indicator" will denote any one or all of the following variables: CEU2023600001, CEU2023700001, AHECONS, HOUSTNSA, UNRATENSA, CEU2000000001, CCI, and BCI. These indicators are employed as explanatory variables within the time series forecasting model. Collectively, the indicators encompass aspects of labor supply and costs, construction cost, employment, and economic strength (HOUSTNSA) parameters. It should be noted that, with the exception of PRRESCON, other variables listed in Table 1 are designated as "target variables". This nomenclature is applied because they are the parameters the model seeks to predict.

3.2. Simple Vector Autoregression (VAR) Model

The study utilizes the standard Augmented Dickey–Fuller (ADF)–Granger causality test procedure [32–34]. Important results include that of the Granger causality test, but since both ADF and the Granger causality test are well-known methodologies, in the interest of space, the VAR model configuration will be emphasized instead.

Two types of VAR models are constructed for the purpose of this paper. The first one is the simple VAR model, constructed using the target variable and only one of its indicators. The purpose of such configuration is to negate the potential effect of multi-collinearity and ensure the maximum interpretability of the model. A simple model is constructed between each target variable (construction spending) and each potential indicator, yielding a total of 72 models (9 target variables × 8 potential indicators). Impulse–response relationships derived from these simple models will elucidate the interpretative value of the model.

The simple models are named as "Simple Model [Target Variable]-[Indicator]". For example, the simple model constructed using PNRESCON and UNRATENSA is designated as Simple Model PNRESCON-UNRATENSA, using the following equation:

$$\mathrm{PNRESCON_t} = \alpha_0 + \alpha_1 \mathrm{PNRESCON_{t-1}} + \cdots + \alpha_p \mathrm{PNRESCON_{t-p}} + \beta_1 \mathrm{UNRATENSA_{t-1}} + \cdots + \beta_p \mathrm{UNRATENSA_{t-p}} + \mathrm{u_t}$$

where the lag order p is determined by AIC [35] as the selection criterion. In brief, the VAR model construes the current value of the target variables to be a linear combination of the p number of past values of the target variables itself and the past value of the leading indicator.

3.3. Aggregate Forecasting Model

The other type of VAR model is the aggregate forecasting model, which incorporates all potential indicators at once for each target variable of construction spending, resulting in 9 models.

For example, the aggregate forecasting model for PNRESCON is:

$$\begin{aligned} \mathrm{PNRESCON_t} &= \alpha_0 + \alpha_1 \mathrm{PNRESCON_{t-1}} + \cdots + \alpha_p \mathrm{PNRESCON_{t-p}} \\ &+ \sum_{indicators} \begin{bmatrix} \beta_1 \mathrm{CEU2023600001_{t-1}} & \cdots & \beta_p \mathrm{CEU2023600001_{t-p}} \\ \vdots & \ddots & \vdots \\ \delta_1 CEU2000000001_{\mathrm{t-1}} & \cdots & \delta_p CEU2000000001_{\mathrm{t-p}} \end{bmatrix} + \mathrm{u_t} \end{aligned}$$

where α, β, δ, etc., refer to the model coefficient, and the lag order p is determined by the best fit under AIC. The aggregate model is, at its core, a multivariate version of the single VAR model.

This model is constructed to be as comprehensive as possible, but potentially disregards the issue of multi-collinearity. The purpose of the configuration is to demonstrate the raw forecasting power of the VAR model and highlight the full predictive capacity of the VAR framework.

4. Results

4.1. Granger Causality Test Results

4.1.1. Total Construction

The Granger causality test was applied across lags of multiples of 3, from 3 to 27. A detailed summary of the Granger causality test results is presented in Table 2. By "leading indicator", we refer to the fluctuation of the past value of the leading indicator being explanatory of the current fluctuation of the target variable. The strength of the correlation is qualitatively construed, using the number of lags at which statistically significant correlation exists, and the strength of individual correlation. For example, CEU2000000001 shows significant bonds with TTLCON over all of the lags observed, while CCI only demonstrates correlation at lag 9; thus, CEU2000000001 is considered to be more strongly correlated with TTLCON compared to CCI. The rationale behind the classifications is that, the higher the number of lags at which statistically significant correlation exists, the more likely a leading indicator will be explanatory of the target variable in a multivariate time-series model (the lag order of which may vary situationally). Additionally, in reference to Ahmadi and Shahandashti 2017 [4], having more lags at which a statistically significant correlation exist also qualitatively means that the leading indicator is explanatory of the target variable over a larger period of time rather than in only a particular timeframe.

Table 2. Granger causality test results between TTLCON and other vital series.

Null Hypothesis	**F-Statistics**								
	Lag 3	**Lag 6**	**Lag 9**	**Lag 12**	**Lag 15**	**Lag 18**	**Lag 21**	**Lag 24**	**Lag 27**
ΔCEU2000000001 is not a leading indicator of ΔTTLCON	15.55 ***	18.79 ***	31.21 ***	4.52 ***	4.91 ***	4.41 ***	3.44 ***	3.23 ***	2.78 ***
ΔCEU2023600001 is not a leading indicator of ΔTTLCON	6.95 ***	11.76 ***	20.29 ***	1.7	2.52 **	2.26 **	1.74 *	2.06 **	1.83 *
ΔCEU2023700001 is not a leading indicator of ΔTTLCON	39.2 ***	28.94 ***	35.0 ***	4.1 ***	4.47 ***	3.81 ***	2.86 ***	2.66 ***	2.58 ***
ΔAHECONS is not a leading indicator of ΔTTLCON	2.02	11.02 ***	9.13 ***	1.29	1.04	1.13	1.01	0.88	0.86
ΔHOUSTNSA is not a leading indicator of ΔTTLCON	8.04 ***	8.72 ***	5.36 ***	0.61	0.76	0.6	0.71	0.61	0.68
ΔCCI is not a leading indicator of ΔTTLCON	0.78	1.88	2.11 *	1.56	1.25	1.2	1.25	1.28	1.41
ΔBCI is not a leading indicator of ΔTTLCON	2.2	1.07	1.46	1.34	1.23	1.03	0.94	1.12	1.12
ΔUNRATENSA is not a leading indicator of ΔTTLCON	11.66 ***	42.68 ***	46.8 ***	2.54 **	2.33 **	2.06 **	1.71 *	1.72 *	1.68 *

* Rejection of null hypothesis at 5% significance level; ** rejection of null hypothesis at 1% significance level; *** rejection of null hypothesis at 0.1% significance level.

The analysis reveals that the total construction spending of the United States (TTLCON) is strongly correlated with employment conditions and moderately correlated with economic strength and labor cost. Contrary to expectations, the correlation between TTLCON and CCI appears weak, and no significant causal relationships were observed between TTLCON and BCI.

4.1.2. Summary of All Granger Causality Test Results

Nine sets of causality test results of the same scale as Table 2 were generated as a result of the study. In the interest of space, these results will not be fully displayed in the paper but will be available upon request. Instead, this paper presents the rest of the causality test results in a summarized correlation table (Table 3), with the causal relationship between target variables with each of the indicators categorized as "None", "Weak", "Moderate", or "Strong".

Table 3. Correlation table of Granger causality test results.

	TTLCON	**PLODGCON**	**PNRESCON**	**PRCOMCON**	**PREDUCON**	**PRHLTHCON**	**PRMFGCON**	**PROFCON**	**PRRELCON**
CEU2000000001	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Strong
CEU2023600001	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Strong
CEU2023700001	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Strong
AHECONS	Moderate	Strong	Moderate	Moderate	Strong	Moderate	Strong	Moderate	Strong
HOUSTNSA	Moderate	Moderate	Moderate	Strong	Moderate	Strong	Strong	Strong	Strong
CCI	Weak	Weak	Moderate	Weak	Weak	Moderate	Weak	None	None
BCI	None	Moderate	Weak	Moderate	Weak	Weak	None	None	Moderate
UNRATENSA	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Strong

For each pair of target variable–indicator, the categorization/scoring scheme has been adopted as follows:

$$Points = 1 \times \#of\ lags\ with\ 5\%\ significance + 2 \times \#of\ lags\ with\ 1\%\ significance$$
$$Point\ of\ 0 = None,\ Points\ of\ 0-3 = Weak,\ Points\ of\ 3-8 = Moderate,\ Points\ of\ 8+ = Strong$$

It should be noted that this scoring scheme is arbitrarily determined for the specific purpose of this study. Nevertheless, it provides a useful qualitative assessment summary of the strength of causality, as summarized in the correlation table (Table 3).

It is observed that all types of construction spending display a consistently strong correlation with employment conditions (CEU2000000001, CEU2023600001, CEU2023700001, and UNRATENSA). Additionally, all types of construction spending observed are at least moderately correlated with economic strength (HOUSTNSA) and construction wage (AHECONS). The correlations between various types of construction spending and BCI/CCI are found to be variable, with correlation strength not exceeding moderate.

4.2. Impulse–Response Function Result Summary

This section presents the notable impulse–response relationships between TTLCON/PNRESCON and each indicator in the following figures and all others in a table format.

4.2.1. Impulse (Indicators)–Response (TTLCON) Functions

Similar to residential construction spending, which is observed in Zhang, Yang, and Wang 2023, TTLCON also exhibits a consistent response to past escalations in construction cost and unemployment rate. Figure 1 delineates the impulse–response relationship between UNRATENSA/CCI and TTLCON under optimal VAR specifications. Similar to those observed in Zhang, Yang, and Wang 2023 [10] concerning PRRESCON, past escalations in UNRATENSA/CCI also precipitate a subsequent decrease/increase in TTLCON. However, the response benchmarks of TTLCON diverge from those associated with PRRESCON. In over 95% of observed instances, an escalation in UNRATENSA at any given time is likely to result in a decrease in TTLCON 6 and 10 months later; similarly, escalations in CCI are found to cause an increase in TTLCON 5 months thereafter.

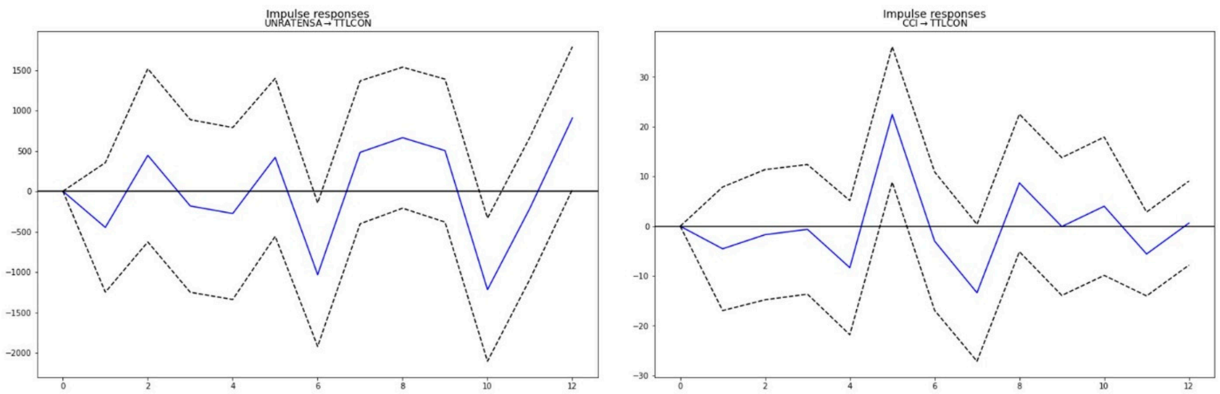

Figure 1. Impulse (UNRATENSA)–response (TTLCON) curve (**Left**); impulse (CCI)–response (TTLCON) curve (**Right**); blue line: mean response magnitude; dashed line: 95% confidence interval.

It is well documented that construction spending serves as an economic stimulus; thus, it is reasonable to assume that TTLCON would be positively correlated with economic strength indicators. The data illustrated in Figure 2 corroborate this hypothesis, revealing that past escalations in HOUSTNSA are positively correlated with subsequent escalations in TTLCON 1 and 5 months later. The negative impulse at a lag of 9 months is counterbalanced by an equivalent positive impulse at a lag of 11 months, rendering the former inconsequential.

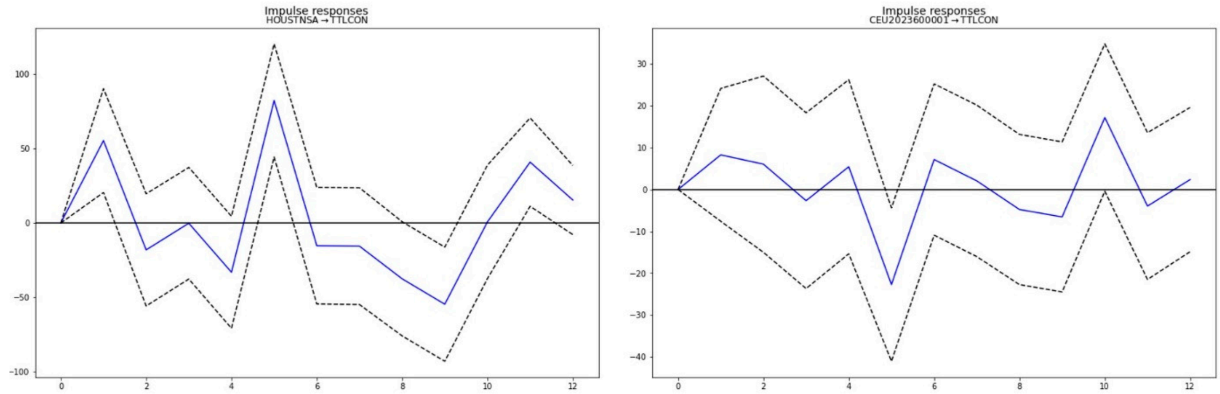

Figure 2. Impulse (HOUSTNSA)–response (TTLCON) curve (**Left**); impulse (CEU2023600001)–response (TTLCON) curve (**Right**); blue line: mean response magnitude; dashed line: 95% confidence interval.

A representative impulse–response relationship between TTLCON and labor supply is provided in Figure 2. Conforming with the findings of Zhang, Yang, and Wang 2023 [10] and Simonson 2013 [9], whose presentation in the second slide shows an approximate 11-month lag between the lowest point of construction employment and total construction spending between 2006 and 2016, in more than 95% of cases, an increase in CEU2023600001 precedes an increase in TTLCON 10 months later. Conversely, a statistical pattern emerges where an escalation in labor supply causes a decrease in TTLCON 5 months later in more than 95% of the cases. A similar dynamic was observed in Zhang et al., 2023 [10] between PRRESCON and construction cost/labor supply. An escalation in supply, likely in response to a higher price, can compromise the short-term feasibility of construction projects.

4.2.2. Impulse (Indicators)–Response (PNRESCON) Functions

Figure 3 summarizes the impulse–response relationships between HOUSTNSA and PNRESCON as well as UNRATENSA and PNRESCON. Similar to TTLCON, PNRESCON has a positive correlation with the overall economic strength of the United States and a negative relationship with unemployment. Regarding the impulse (UNRATENSA)–response (PNRESCON) function, the TTLCON's benchmarks for negative response to the impulse unemployment rate, namely, lag 6 and lag 10, remain valid for PNRESCON.

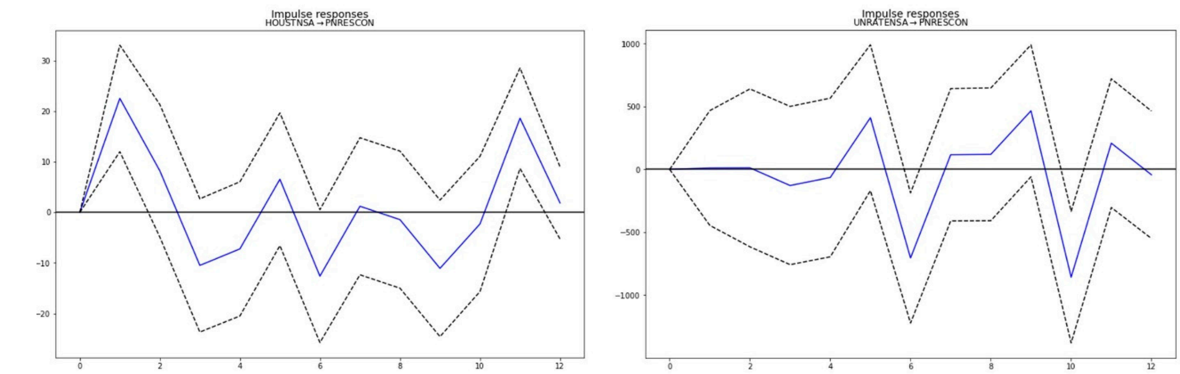

Figure 3. Impulse (HOUSTNSA)–response (PNRESCON) curve (**Left**); impulse (UNRATENSA)–response (PNRESCON) curve (**Right**); blue line: mean response magnitude; dashed line: 95% confidence interval.

Figure 4 summarizes the impulse–response relationships between CEU2023600001 and PNRESCON as well as AHECONS and PNRESCON. Multiple statistically significant impulse–response relationships at the 95% confidence level are observed between CEU2023600001 and PNRESCON, but a negative response at a 6-month lag is offset by a subsequent positive response of similar magnitude at a 7-month lag immediately after. Ultimately, only the positive responses of PNRESCON to CEU2023600001 at a lag of 10 and 12 months are notably significant. Additionally, the impulse–response function between AHECONS and PNRESCON reflects a time-varying effect of labor wage escalation on construction spending. In more than 95% of the cases, an increase in AHECONS was observed to have an immediate negative impact on PNRESCON 3–4 months later but a positive impact 7 months later. PNRESCON's reactions to past shifts in CEU2023600001 and AHECONS are consistent with the findings presented in Zhang, Yang, and Wang 2023 [10].

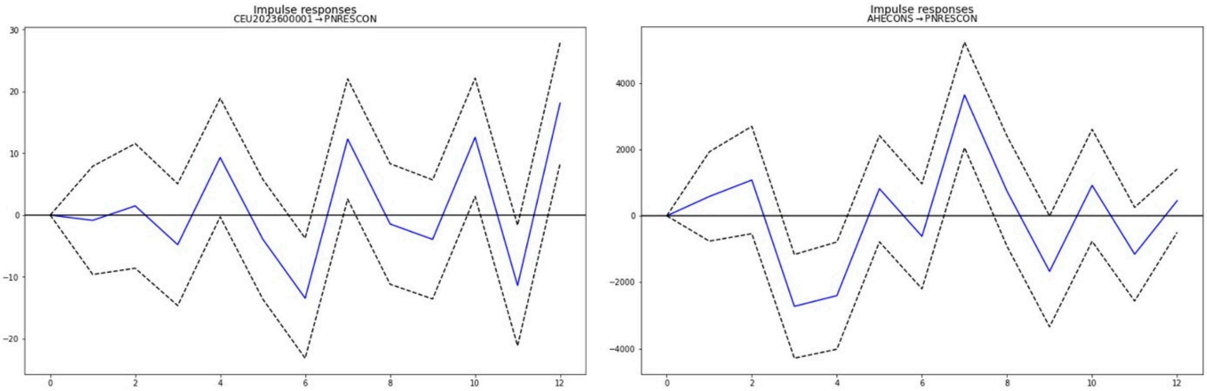

Figure 4. Impulse (CEU2023600001)–response (PNRESCON) curve (**Left**); impulse (AHECONS)–response (PNRESCON) curve (**Right**); blue line: mean response magnitude; dashed line: 95% confidence interval.

A summary of significant impulse (indicators)–response (types of construction spending) relationships, including their signs and the lag at which they occur, is presented in Table 4. The figures refer to the time lag of the responses (in number of months), and the signs refer to the sign of the responses with the purpose of observing whether the correlations are positive or negative. The analysis addresses the knowledge gap defined in the introductory section by adding model interpretation on top of causality test results. More extensive discussion of these relationships is provided in subsequent sections.

Table 4. Table summary of 95%+ significant impulse–response relationships, with figures indicating the time-lagged-ness (in months) and signs indicating the signs of the responses towards impulse of impulse indicators (convention: +x—positive response of the response variables to impulse indicators escalations at a lag of x months; −y—negative response of the response variables to impulse indicators escalations at a lag of y months).

	Impulse Indicators							
Response Variables	**HOUSTNSA**	**CEU2023600001**	**CEU2023700001**	**CEU2000000001**	**AHECONS**	**CCI**	**BCI**	**UNRATENSA**
TTLCON	+1 +5	−5 +10	−5 +7 +11	−5 +10	N/A	+5	N/A	−6 −10
PNRESCON	+1 +11	+10 +12	N/A	N/A	−3 −4 +7	−3 +5	+5	−6 −10
PLODGCON	N/A	N/A	N/A	N/A	+5 −6 +7	N/A	N/A	N/A
PRCOMCON	N/A	N/A	+10	+10	−3 +6 +7	+10	N/A	−6 −10
PREDUCON	+1	+10 +12	N/A	+12	+7 +8 −9 +10	−4	−4	N/A
PRHLTHCON	N/A	N/A	N/A	+10	−3 −4 +7	−6	−6	N/A
PRMFGCON	+5 +11	N/A	N/A	N/A	N/A	N/A	N/A	N/A
PROFCON	N/A	+10	N/A	N/A	N/A	N/A	N/A	N/A
PRRELCON	N/A	N/A	N/A	−7	N/A	−7	−7	N/A

4.3. Forecasting Results

The forecasting results for the target variables (various categories of construction spending) during the testing period of M1 2018 to M12 2022 are summarized in this section. Three representative cases of PNRESCON, TTLCON, and PLODGCON are visually presented, while results for additional categories are summarized in a table format.

The forecasting results for TTLCON, PNRESCON, and PLODGCON are depicted in Figure 5. The COVID-19 pandemic caused major disruptions in the economy, resulting in the disbanding of some of the established regression-based forecasting models prior to the pandemic [22,36]. The forecasting results from this paper show that the disruptions in construction spending induced by the pandemic vary distinctively by construction type. For the three cases presented, first, TTLCON has witnessed a surge in spending, outpacing the forecasts of regression models. This trend corroborates with the findings of Zhang, Yang, and Wang 2023 [10], who discovered that residential construction spending (which makes up roughly two-thirds of TTLCON) increased dramatically as a result. Second, PNRESCON does not show significant deviations from regression-based model predictions. Lastly, on the other hand, PLODGCON experienced an atypical decline in spending, suggesting a sector-specific variability to the pandemic's impacts.

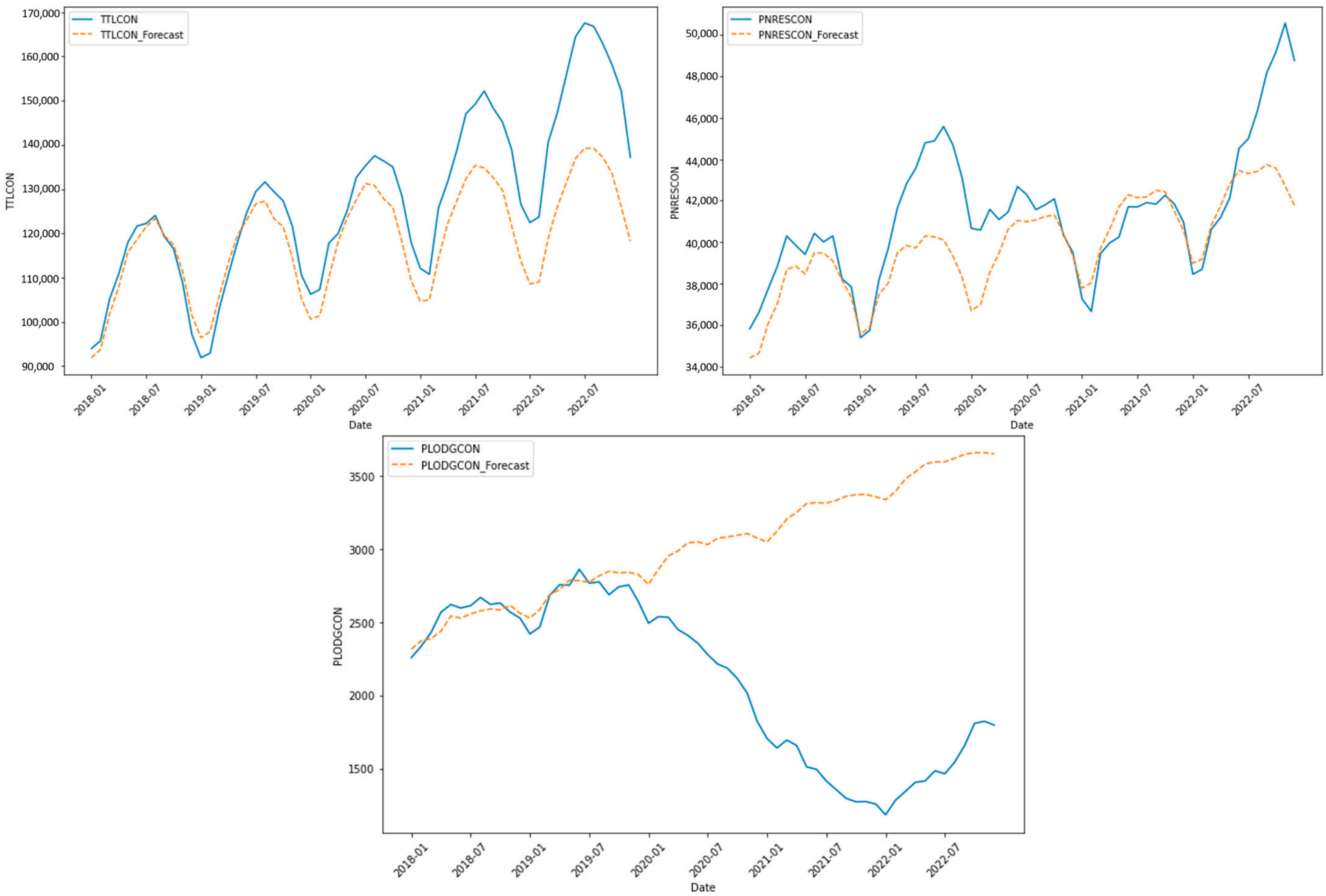

Figure 5. Forecasting results for TTLCON (**Upper Left**), PNRESCON (**Upper Right**), and PLODGCON (**Lower Middle**).

Forecasting results for additional categories of construction spending are summarized in Table 5. Two Mean Absolute Percent Error (*MAPE*) figures for the testing periods of M1

2018–M12 2019 and M1 2018–M12 2022 were calculated for each construction spending type, using the following equation:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} \frac{|T_t - F_t|}{T_t} \times 100\%$$

where T_t refers to the testing data at a time, F_t refers to the forecasted data at that time in the testing period, and n refers to the number of observations during the testing period.

Table 5. Summary of forecasting results.

Construction Spending Type	MAPE (%) M1 2018–M12 2019	MAPE (%) M1 2018–M12 2022	Description of COVID-19 Disruption
TTLCON	2.79	6.99	Escalation
PNRESCON	4.94	4.15	Steady
PLODGCON	2.68	60.9	De-escalation
PRCOMCON	13.6	12.6	Steady
PREDUCON	6.25	8.69	De-escalation
PRHLTHCON	3.66	6.86	Escalation
PRMFGCON	6.40	8.85	Escalation
PROFCON	9.56	7.85	Steady
PRRELCON	6.53	12.4	De-escalation

These MAPE figures reveal that, during the first two years of the testing period, the forecasts are decently accurate, meaning that the models constructed using pre-COVID-19 data are adequate by pre-COVID-19 standards. However, when the testing period observed stretches to M1 2018–M12 2022, the average error drastically increases for most construction spending types as a response.

Visual aids illustrating these forecasting results can be provided upon reasonable request to the corresponding author. Generally speaking, the raw values of the increased error during the latter three years of the testing period show consistent monthly trends. These consistent error trends provide qualitatively describable COVID-19 disruptions, as summarized in the column "Description of COVID-19 Disruption". The observation that the COVID-19 pandemic introduced anomalies during the testing period does not undermine the validity of the findings presented in this paper. Drawing from the principles outlined by Perron 1989 [37], it can be argued that shocks typically have a temporary, transient impact, and the core conclusions of this study are anticipated to retain their reliability over the long-term, despite the short-term COVID-19 disruptions.

5. Discussion and Lessons Learned

5.1. Interpreting the Model

Up to now, the main body of forecasting literature in the field of construction that employs regression-based time-series models relies mainly on comparing the forecasting results with the testing data and then computing the average error for result verification purposes [2,3,28,29]. Such an approach is correct, as the accuracy of the forecast is one of the most important factors in the determination of a model's value.

Nevertheless, simply focusing on forecasting itself also means missing out on much empirical evidence that could be generated through model interpretation, which is key for inferring causality within correlated phenomena. In this vein, the discussions below highlight the importance of model interpretation. Multiple impulse–response trials were carried out between type-specific construction spendings and their potential indicators, yielding results that conform with existing well-substantiated evidence.

The most prominent example of such consistency lies in the positive response of present TTLCON and PNRESCON to the increase in labor supply 10–12 months in the past. Such an impulse–response relationship directly conforms with Simonson's 2013 [9] findings and intuitively aligns with the construction industry's operational dynamics, where the occurrence of a building project does not generate jobs; rather, it is the process of construction that does. Therefore, it is reasonable to expect a time-lagged correlation between construction spending (calculated at the time when projects are put in place) and labor supply, the lag of which corresponds to the length of time projects typically get built in.

Another example of such consistency lies in the positive time-lagged response of TTLCON/PNRESCON towards past escalations of economic strength (HOUSTNSA), and the negative time-lagged response of construction spendings to UNRATENSA. The results conform with the general belief that construction spending is positively correlated with economic conditions, alongside the time-lagged correlation discovered in Ahmadi and Shahandashti 2017 [4].

TTLCON and PNRESCON are also observed to be positively correlated with an indicator of economic strength (HOUSTNSA) and negatively correlated with the unemployment rate, conforming to the findings of the literature that draws a positive correlation between construction and employment level. Construction cost was observed to be positively correlated with TTLCON, which conforms to the fact that construction cost makes up much of the value for residential development and stimulates PRRESCON [10], which takes up roughly two-thirds of TTLCON. PREDUCON and PRRELCON were shown to have a time-lagged negative correlation with escalation in construction cost, conforming with Fernández et al.'s 2023 [23] discussion on resource limitations in educational construction and the financial constraints in non-profit religious developments whose feasibility is compromised as a result of cost escalation.

Interesting empirical evidence derived using statistical analysis is the time-varying impact of labor supply and labor cost, as evident in the response of PNRESCON, PLODGCON, PRCOMCON, PREDUCON, and PRHLTHCON towards a past escalation of AHECONS. The immediate impact of AHECONS at a lag of 3–4 months is observed to be negative, but at a lag of 7–10 months, it becomes positive.

5.2. Analyzing as Type-Specifically as Possible

It was hypothesized that different types of construction spending will react differently to the same suite of indicators. Our analysis shows that such appears to be the case.

When resolving similar problems, researchers must distinguish between various types of construction spending, as this case study presents an instance in which type-specific construction spending exhibits varying reactions to factors in causality tests, model interpretation, and forecasting. From the analysis, it can be observed that different kinds of type-specific construction spending are not well correlated with each other. And the results applicable to one type of construction spending may not be valid for another. This diversity is reflected in the varying significance of the Granger causality tests, contradictory impulse–response relationships, and the myriad ways in which forecasts are disrupted by the COVID-19 pandemic.

The type of construction spending may influence the existence of causal relationships. Observing the Granger causality test results outlined in this case study, it is evident that not all indicators correlate with construction cost, as TTLCON and PRMFGCON correlate only with CCI, while PRRELCON correlates only with BCI. There are even cases, such as PROFCON, that are not correlated with CCI at all. On top of all that, PRRESCON was observed to be strongly correlated with both BCI and CCI [10], a trait that is not shared with any other construction spending type.

As stated in the previous sections, impulse–response analysis indicates different types of construction spending display different time-lagged responses to the indicators. The most important case in this case study is the response construction cost; while PREDUCON, PRHLTHCON, and PRRELCON react negatively to past construction cost escalations, TTLCON reacts positively.

Lastly, the time series for different kinds of type-specific construction spending react differently to disruptions during the testing period, as described in the forecasting results section.

5.3. Using Complex Models for Forecasting

This case study employs an aggregate VAR model that encompasses all potential indicators simultaneously when conducting the forecast, without addressing potential multi-collinearity issues. The purpose of such an approach is to demonstrate that complex models, which incorporate multiple factors, tend to exhibit strong forecasting capacity. Okuta et al., 2023 [38] presented a representative case where a complex time-series model was deemed more accurate than simpler models when forecasting housing prices in Kenya, and such trait is also applicable to construction spending forecasting.

Figure 6 demonstrates a case where the aggregate model significantly outperformed a simpler model, which only used PNRESCON and UNRATENSA, in forecasting PNRESCON. Generally, the simpler the model, i.e., the fewer the regression features, the more forecasting variance would have to be explained using the past value of target variable itself. Therefore, when the target variable displays inconsistent statistical properties, such as varying seasonality patterns or frivolous trends, the forecasting capabilities of simpler models will likely be compromised. This concern, though not explicitly addressed in past work, does reflect itself in the existing literature. Abiniangerabi et al., 2017 [2] forecasted the aggregate construction spending (such as residential) of a large area (United States). Their construction spending sample size was large and displayed steady statistical properties, and thus the paper could make do with a simple model constructed with only construction spending and ABI. On the other hand, Sing et al., 2015 focused on a smaller sample size (construction of HK), and their model had to utilize a moderately more complex model of three to four indicators [39].

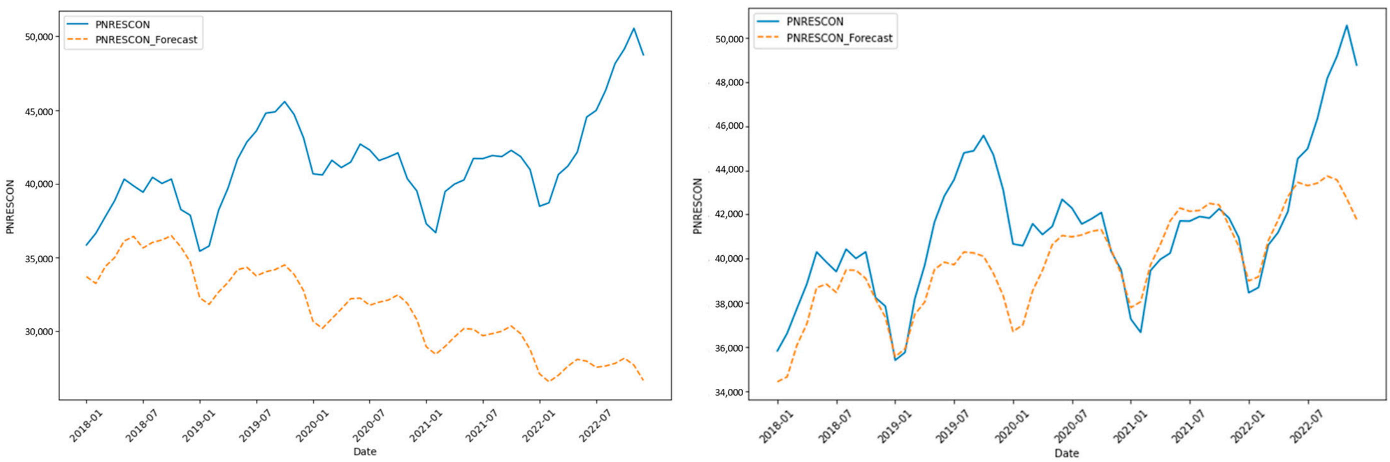

Figure 6. Forecast result of simple model for PNRESCON-UNRATENSA (**Left**), and forecast result of aggregate forecasting model for PNRESCON (**Right**).

However, the fact that the simple model PNRESCON-UNRATENSA did a suboptimal job at forecasting the target parameter does not mean that the model was incorrectly constructed. Instead, it indicates that some characteristics of PNRESCON during the testing period could only be explained by indicators other than UNRATENSA or past values of PNRESCON.

6. Conclusions

Total and type-specific construction spending were observed in this study for correlation and forecasting performance. Strong correlations were universally observed between construction spendings and labor supply/unemployment/economic output.

However, the different kinds of type-specific construction spending display varying behaviors that should not be generalized. For one, they display a different level of Granger correlation and impulse–response relationship with construction cost. For another, different construction spending types react differently to the COVID-19 disruption, with some escalating as a result of the pandemic, some de-escalating, and others, such as PNRESCON, remaining steady. The diversity in response means that the economic analysis on construction should be as type-specific as possible.

This paper also provides methodological advice for forecasting studies. The current body of literature conducts result verification by comparing forecast with testing data. This study demonstrates that, through model interpretation, correlation can be used to infer causality. Impulse–response functions revealed a time-lagged positive correlation between construction spending and escalation in labor supply, conforming to findings of relevant studies and the operational dynamics of construction; and a time-lagged positive relationship between construction spending and economic strength indicators, conforming to Ahmadi and Shahandashti 2017 [4], who found economic strength to be indicative of construction spending.

The policy implication of this paper is three-fold: for one, although this paper finds a significant correlation between construction spending and employment, we argue that construction spending is not a proper leading indicator for employment, as the completion of a building project does not generate jobs; rather, it is the process of construction that does. Therefore, construction spending tends to lag (occur after) employment conditions rather than being leadingly indicative of them.

For another, during COVID-19, the United States saw a surge in construction, which causes many to speculate that disruptions in relevant time series in the form of an unusual escalation have occurred. This is also not the case, as depending on the construction type, the behavior of construction spending can either be an escalation, a de-escalation, or remain steady without shocks during COVID-19.

Lastly, what should be exclusively mentioned is that construction cost (BCI/CCI) needs to be factored in in moderation when constructing models. Granger trials showed only weak to moderate correlation between various types of construction spending. Labor supply/construction demand carries considerably more explanatory power.

The next steps in research lie in the exploration of relationships between construction spending and economic indicators with "noisy" temporal dynamics (it was shown in Table 1 that most data used have seasonality patterns), with a focus on long-term relationships that do not attenuate over time. It was shown in Table 4 that some of the impulse variables including HOUSTNSA have a time-varying impact on the future value of construction spending. This could either be attributable to a unique but consistent pattern, by which said indicators interact with construction spending, but it could also be that the temporal association observed by the VAR model is overly fuzzy, and that it is more worthwhile to explore said relationships using an error correction model instead. We will leave it to future research.

Author Contributions: Conceptualization, X.Z.; Methodology, X.Z.; software, Y.W. and S.X.; validation, E.Y.; formal analysis, X.Z.,Y.W. and S.X.; investigation, Y.W.; resources, L.M.; data curation, X.Z. and S.X.; writing—original draft preparation, X.Z.; writing—review and editing, E.Y.; visualization, Y.W.; supervision, X.Z. and E.Y.; project administration, X.Z. and L.M.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data Used in this research can be downloaded from https://fred.stlouisfed.org/ (accessed on 1 December 2023) and from Engineering News Record with a subscription.

Conflicts of Interest: Xingrui Zhang was employed by China Construction Infrastructure Corp., a branch of China State Construction Engineering Corp. Yunpeng Wang was employed by the 2nd Engineering Bureau, China State Construction Engineering Corp. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Davis, M.A.; Heathcote, J. Housing and the business cycle. *Int. Econ. Rev.* **2005**, *46*, 751–784. [CrossRef]
2. Abediniangerabi, B.; Shahandashti, S.M.; Ahmadi, N.; Ashuri, B. Empirical Investigation of Temporal Association between Architecture Billings Index and Construction Spending Using Time-Series Methods. *J. Constr. Eng. Manag.* **2017**, *143*, 04017080. [CrossRef]
3. Ashuri, B.; Shahandashti, S.; Lu, J. Empirical test for identifying leading indicators of ENR construction cost index. *Constr. Manage. Econ.* **2012**, *30*, 917–927. [CrossRef]
4. Ahmadi, N.; Shahandashti, M. Comparative empirical analysis of temporal relationships between construction investment and economic growth in the United States. *Constr. Econ. Build.* **2017**, *17*, 85–108. [CrossRef]
5. Landers, J. Increased spending on US infrastructure would boost economy, report says. *Civ. Eng. ASCE* **2014**, *84*, 14–15.
6. Turin, D.A. *The Construction Industry: Its Economic Significance and Its Role in Development*; Environmental Research Group: London, UK, 1969.
7. Ball, M.; Wood, A. How many jobs does construction expenditure generate? *Constr. Manag. Econ.* **1995**, *13*, 307–318. [CrossRef]
8. Hassan, M.E.M. Modeling the Macroeconomic Impact of Construction Spending as a Public Policy Tool—The Case of Employment. Ph.D. Thesis, Purdue University, West Lafayette, IN, USA, 2017.
9. Simonson, K. Construction Spending, Labor & Materials Outlook. *Change* **2013**, *15*, 30.
10. Zhang, X.; Yang, E.; Wang, Y. Time series observation of relationship between United States private residential construction spending and its indicators. *Int. J. Hous. Mark. Anal.* 2023, *ahead-of-print*. [CrossRef]
11. Aaronson, D.; Brave, S.; Cole, R. Using private sector 'big data 'as an economic indicator: The case of construction spending. *Chic. Fed Lett.* **2016**, *366*. [CrossRef]
12. Baker, K.; Saltes, D. Architecture billings as a leading indicator of construction. *Bus. Econ.* **2005**, *40*, 67–73. [CrossRef]
13. Byun, K.J. The US housing bubble and bust: Impacts on employment. *Mon. Lab. Rev.* **2010**, *133*, 3.
14. Aschauer, D.A. Is public expenditure productive? *J. Monet. Econ.* **1989**, *23*, 177–200. [CrossRef]
15. Ofori, G. Construction industry and economic growth in Singapore. *Constr. Manag. Econ.* **1988**, *6*, 57–70. [CrossRef]
16. Tse, R.Y.; Ganesan, S., IV. Causal relationship between construction flows and GDP: Evidence from Hong Kong. *Constr. Manag. Econ.* **1997**, *15*, 371–376. [CrossRef]
17. Gyourko, J.; Saiz, A. Construction costs and the supply of housing structure. *J. Reg. Sci.* **2006**, *46*, 661–680. [CrossRef]
18. Guan, Y.; Cheung, K.S. The Costs of Construction and Housing Prices: A Full-Cost Pricing or Tendering Theory? *Buildings* **2023**, *13*, 1877. [CrossRef]
19. Jud, G.D.; Winkler, D.T. The dynamics of metropolitan housing prices. *J. Real Estate Res.* **2002**, *23*, 29–45.
20. Case, K.E.; Mayer, C.J. Housing price dynamics within a metropolitan area. *Reg. Sci. Urban Econ.* **1996**, *26*, 387–407. [CrossRef]
21. Irandoust, M. House prices and unemployment: An empirical analysis of causality. *Int. J. Hous. Mark. Anal.* **2019**, *12*, 148–164. [CrossRef]
22. Zhang, X.; Yang, E. Have housing value indicators changed during COVID? Housing value prediction based on unemployment, construction spending, and housing consumer price index. *Int. J. Hous. Mark. Anal.* 2023, *ahead-of-print*. [CrossRef]
23. Fernández, R.; Correal, J.F.; D'Ayala, D.; Medaglia, A.L. A decision-making framework for school infrastructure improvement programs. *Struct. Infrastruct. Eng.* **2023**, 1–20. [CrossRef]
24. Alsharef, A.; Banerjee, S.; Uddin, S.J.; Albert, A.; Jaselskis, E. Early impacts of the COVID-19 pandemic on the United States construction industry. *Int. J. Environ. Res. Public Health* **2021**, *18*, 1559. [CrossRef]
25. Araya, F.; Poblete, P.; Salazar, L.A.; Sánchez, O.; Sierra-Varela, L.; Filun, Á. Exploring the Influence of Construction Companies Characteristics on Their Response to the COVID-19 Pandemic in the Chilean Context. *Sustainability* **2024**, *16*, 3417. [CrossRef]
26. Umar, T. The impact of COVID-19 on the GCC construction industry. *Int. J. Serv. Sci. Manag. Eng. Technol. (IJSSMET)* **2022**, *13*, 1–17. [CrossRef]
27. De Henau, J.; Himmelweit, S. A care-led recovery from COVID-19: Investing in high-quality care to stimulate and rebalance the economy. *Fem. Econ.* **2021**, *27*, 453–469. [CrossRef]
28. Ashuri, B.; Lu, J. Time series analysis of ENR construction cost index. *J. Constr. Eng. Manag.* **2010**, *136*, 1227–1237. [CrossRef]
29. Shahandashti, S.M.; Ashuri, B. Forecasting engineering news-record construction cost index using multivariate time series models. *J. Constr. Eng. Manag.* **2013**, *139*, 1237–1243. [CrossRef]
30. Oshodi, O.; Ejohwomu, O.A.; Famakin, I.O.; Cortez, P. Comparing univariate techniques for tender price index forecasting: Box-Jenkins and neural network model. *Constr. Econ. Build.* **2017**, *17*, 109–123. [CrossRef]
31. Williams, T.P. Predicting changes in construction cost indexes using neural networks. *J. Constr. Eng. Manag.* **1994**, *120*, 306–320. [CrossRef]

32. Dickey, D.A.; Fuller, W.A. Distribution of the estimators for autoregressive time series with a unit root. *J. Am. Stat. Soc.* **1979**, *75*, 427–431.
33. Granger, C. Investigating causal relations by economic models and cross-spectral methods. *Econometrica* **1969**, *37*, 424–438. [CrossRef]
34. Said, S.E.; Dickey, D.A. Testing for unit roots in autoregressive moving average models of unknown order. *Biometrika* **1984**, *71*, 599–607. [CrossRef]
35. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [CrossRef]
36. Bhat, M.R.; Jiao, J.; Azimian, A. The impact of COVID-19 on home value in major Texas cities. *Int. J. Hous. Mark. Anal.* **2023**, *16*, 616–627. [CrossRef]
37. Perron, P. The great crash, the oil price shock, and the unit root hypothesis. *Econom. J. Econom. Soc.* **1989**, *57*, 1361–1401. [CrossRef]
38. Okuta, F.O.; Kivaa, T.; Kieti, R.; Okaka, J.O. Comparing simple and complex regression models in forecasting housing price: Case study from Kenya. *Int. J. Hous. Mark. Anal.* 2023, *ahead-of-print*. [CrossRef]
39. Sing, M.C.; Edwards, D.J.; Liu, H.J.; Love, P.E. Forecasting private-sector construction works: VAR model using economic indicators. *J. Constr. Eng. Manag.* **2015**, *141*, 04015037. [CrossRef]