Watt's the Deal? Investigating the role of Virginia's Data Center Industry on US Electricity Markets

by

Antoni Dziwura
Supervisor: Dr Eleanya Nduka

University of Warwick
April 2025

Abstract

The US data centre industry remains the infrastructural backbone of the contemporary Internet. The recent cloud services expansion and 'AI Boom' has highlighted the importance of the industry and resulted in substantial demand for its services (EPRI, 2024). This paper seeks to contribute to the existing literature by analysing the impact of data centres on broader electricity markets. It proposes a three-step strategy: structural break tests (Chow, 1960), (Andrews, 2003) and synthetic control analyses (Abadie and Gardeazabal, 2003) are performed to demonstrate the causal link between data centres and electricity consumption. Vector Autoregressive Models (Sims, 1980) are used to estimate the impact of data centres on regional and sectoral electricity prices and power demand. The study leverages the unique properties of the US state of Virginia and provides quantitative evidence that the data centre industry is a major energy consumer in the Northeastern US.

Keywords: Data Centres, Structural Breaks, Virginia, Vector Autoregression.

JEL Classification: Q41, Q43, Q47.

Word Count: 4900 / 5000

Acknowledgments:

I dedicate this short piece of work first and foremost to my family: Mum, Dad and Stefan, who made me who I am, and to my dear friends: Josh, Jan, Daniel and many more, who motivate me every day to seek the deeper truths. I would like to thank the Trzy Traby Foundation, which enabled my studies through its scholarship, and Professor Robert Akerlof who helped me throughout all my years at Warwick. Finally, shout-out to my peoples, Ali G and J.D. Wetherspoons, who made me love England.

Disclosure

The research was carried out using the R software package. The research incorporated LLM tools offered by OpenAI (2025) and Anthropic (2025) to optimise the underlying R-code, including: bug fixing, graph plotting, guidance on specific libraries or bibliography formatting.

Contents

1 Introduction

Virginia dominates the United States' data centre industry with almost half of all national capacity being installed within the state's borders (JLARC, 2024). The sheer scale of the existing computing infrastructure, along with unique geographic properties of Virginia, offers an opportunity to analyse the impact of data centres on local and regional electricity markets. The study addresses the existing literature gap by deploying previously unused methodologies to provide evidence to what extent data centres influence electricity markets: particularly demand and price patterns. The research aims to provide comprehensive understanding and reliable estimates for future policymakers, ensuring an economically equal and swift growth of the Internet of Things economy.

Section 1 presents the background information and summarises the historical development of the industry. Sections 2 and 3 review the existing literature, methodology and data. Section 4 discusses the results and limitations. I conclude by providing estimated impact of the data centre industry on Virginia's electricity prices, and spillover effects to neighbouring states and sectors.

1.1 Background Information

What is a data centre? A data centre is an infrastructure that *organizes, processes, stores, and disseminates large amounts of data* (Maryland DoC, 2020). Such infrastructure requires substantial amount of resources, primarily electric power, to sustain its services. In 2023, the industry was responsible for 4% of total US national electricity consumption (EPRI, 2024). Geographic location i.e. proximity to major internet cables responsible for the physical data flow of the Internet, and access to reliable power sources, are the defining feature for any computing infrastructure project from an investment perspective (Turner, 2024). Therefore, data centres tend to be located in geographic clusters based on optimal access to internet fibre networks, which are aggregated in *cable landing stations* (IEA, 2024). Virginia has a globally unique, direct access to major subsea internet fibre cables, through e.g., the Virginia Beach landing station, connecting the Northeastern US with EMEA and South America. It also possess a dense on-ground power grid network (EIA, 2024a). As a result, majority of historically US-built data centres were located primarily in Northern Virginia known as the 'Data Center Valley' (Fig. 18). Today, the region has the highest concentration of data centres in the US and globally (Fig. 2).

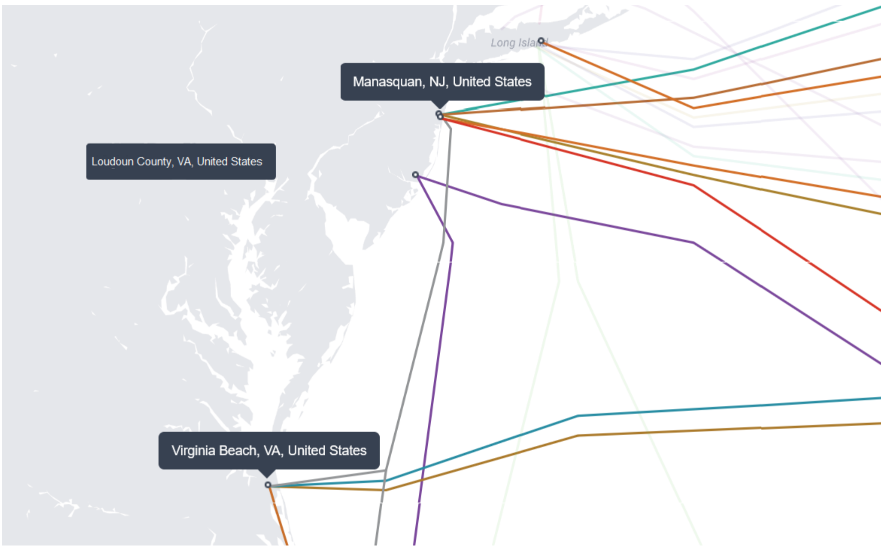
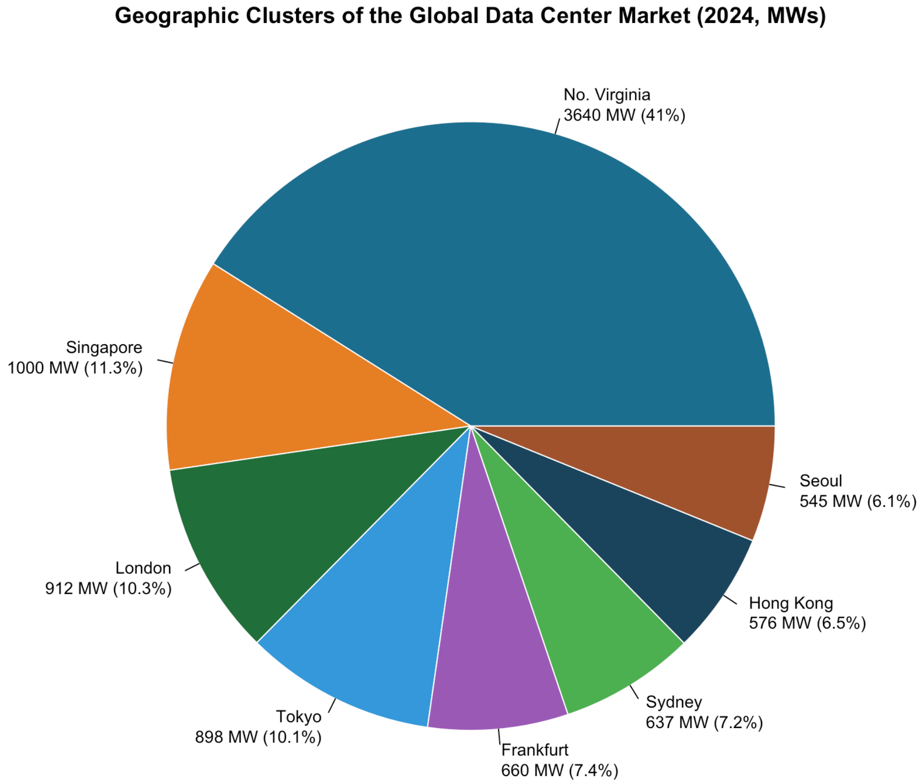


Figure 1: Loudoun County in Northeastern Virginia has the closest proximity to both US-EMEA and US-SA submarine cable infrastructure. Map Source: (TeleGeography, 2025)



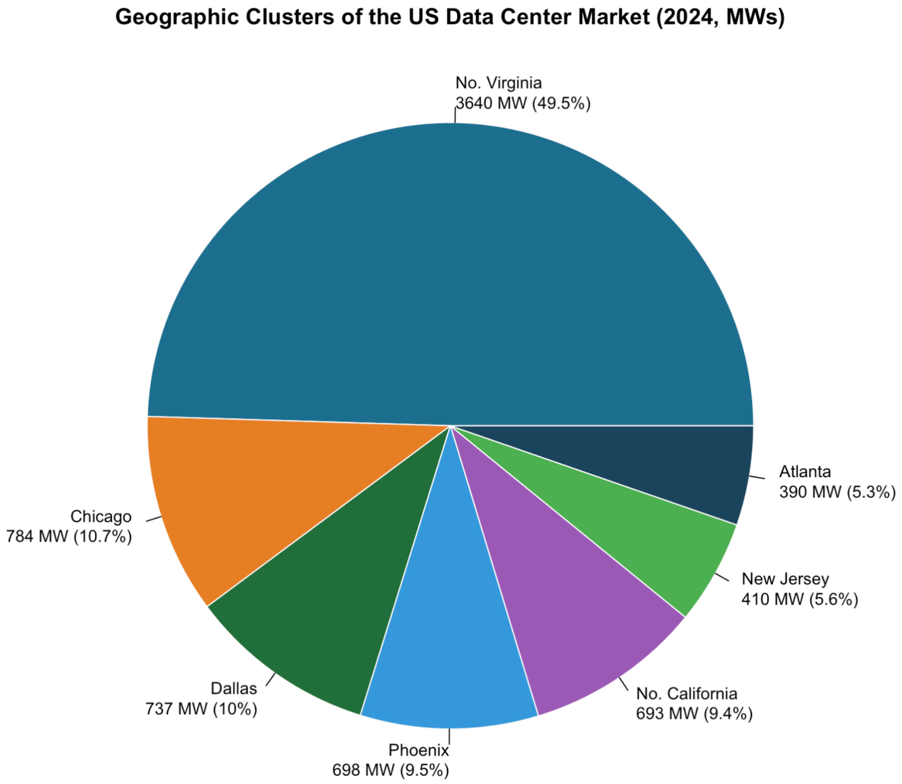


Figure 2: Northern Virginia accounts for almost half of the existing data centres in the US and dominates the global computing industry. Data Source: (Dominion Energy, 2025).

Virginia has massively benefited from its unique geographic location. In recent years, the state has experienced a rapid buildout of new computing infrastructure with annual added capacities of 800 MW: an equivalent of 200 000 households or 45 mid-sized automobile assembly plants

(JLARC, 2024). This resulted in substantial investment inflows; already in 2021 62% of all new investments came from the computing industry (EIA, 2024a), while Loudoun County - the main area of the buildouts, is the richest county *per capita* in the United States (U.S. Census Bureau, 2022).

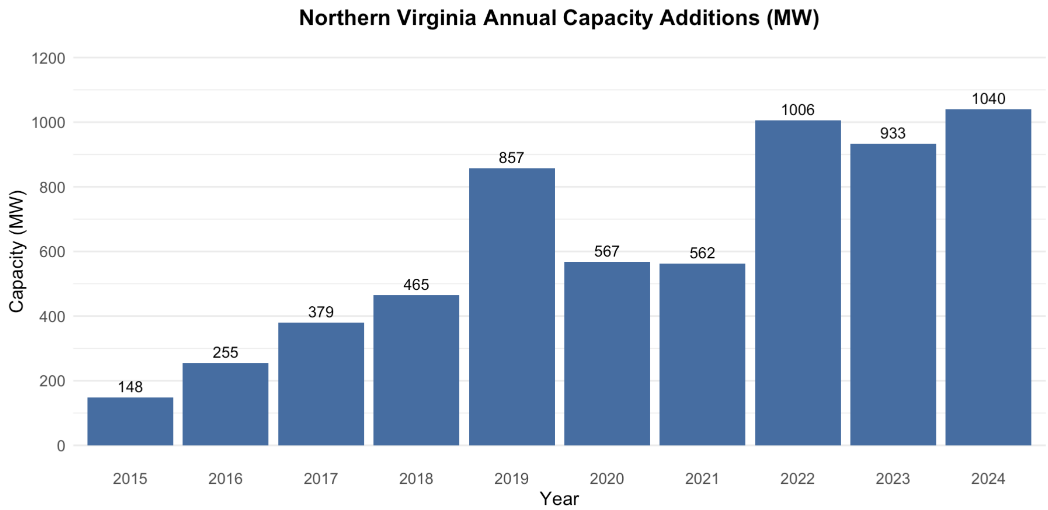


Figure 3: Number of annually installed capacity additions of data centres in Virginia. Data Source: (Dominion Energy, 2025).

The growth of the industry has resulted in unprecedented power demand needed to sustain the newly added infrastructure. Since pre-Covid 2019, Virginia has increased its commercial electricity consumption almost as much as Texas - a state with 3.5 times bigger population, and 4.5 greater GDP (BEA, 2025). Data centres are a major reason for such rapid power demand. They are estimated to account for 25% of Virginia's total electricity consumption in 2023 (EPRI, 2024). As a result, since 2021 the broad commercial sector - to which data centres belong - has overtaken the residential sector as the main power consumer with over a 20% gap in 2024 - a unique phenomenon across US states (See Fig 19). The study seeks to verify whether such power demand shock has impacted local electricity markets.

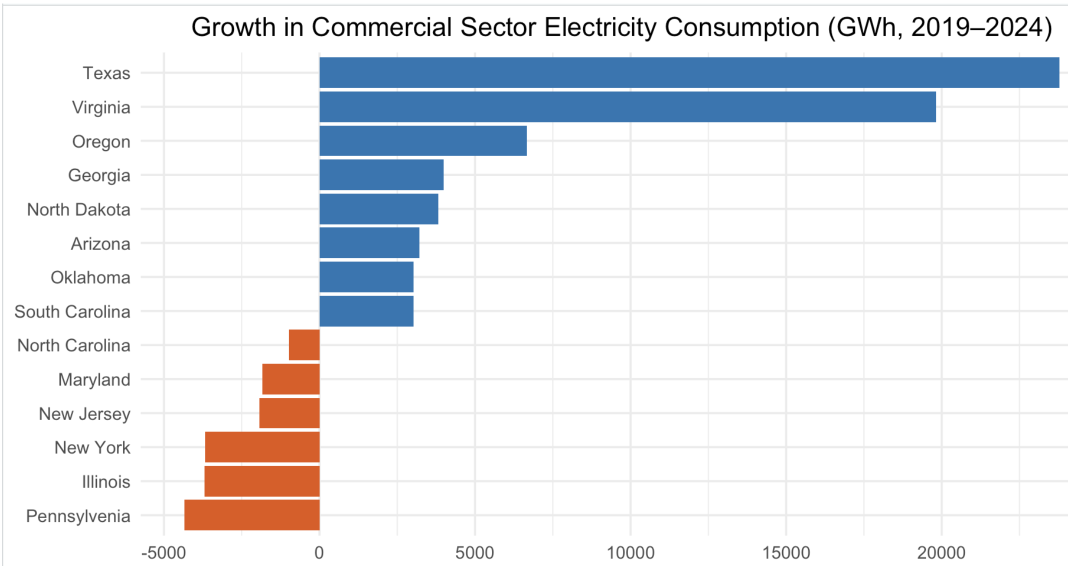


Figure 4: Growth in Commercial Electricity Consumption across US. Data Source: (IEA, 2024)

Importantly for this study, Virginia is the only state in Northeastern US to experience a major growth in commercial electricity consumption, while other regional states do not demonstrate similar patterns of demand (Fig 5 and 6). As a result, the state-based power providers are unable to supply all the locally demanded power: since 2023, Virginia has been the biggest net importer of electricity in the entire United States, importing 36% of state's total power usage (Fig 21) (EIA, 2024). Such massive interstate electricity flows open an opportunity to investigate possible spillover effects across the regional power markets.

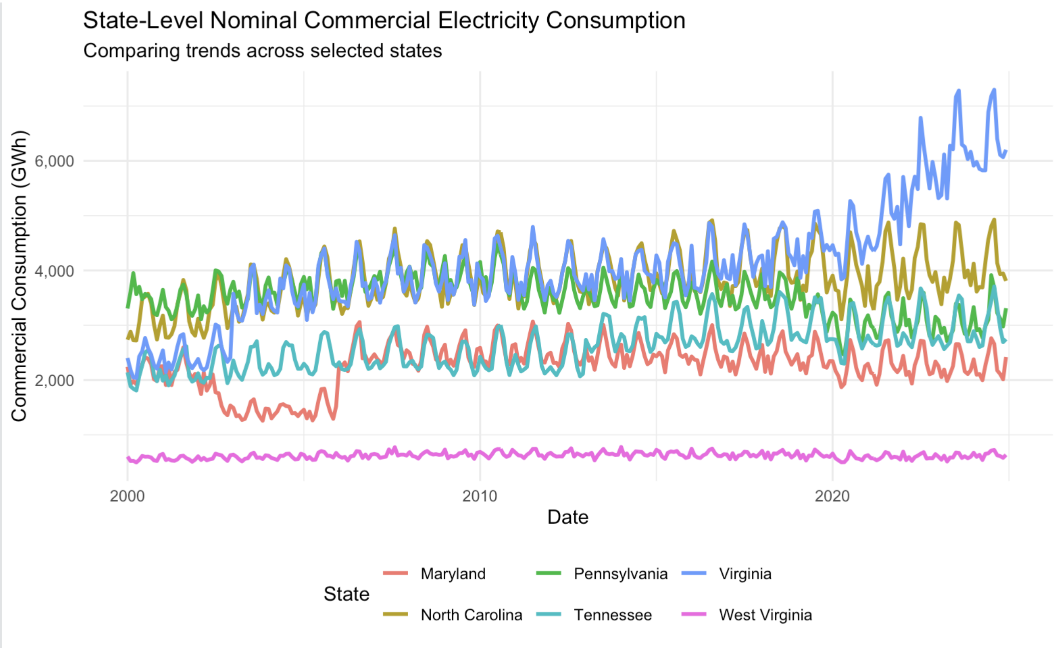


Figure 5: Nominal growth of electricity commercial consumption across Northeastern States. Data Source: (EIA, 2025a)

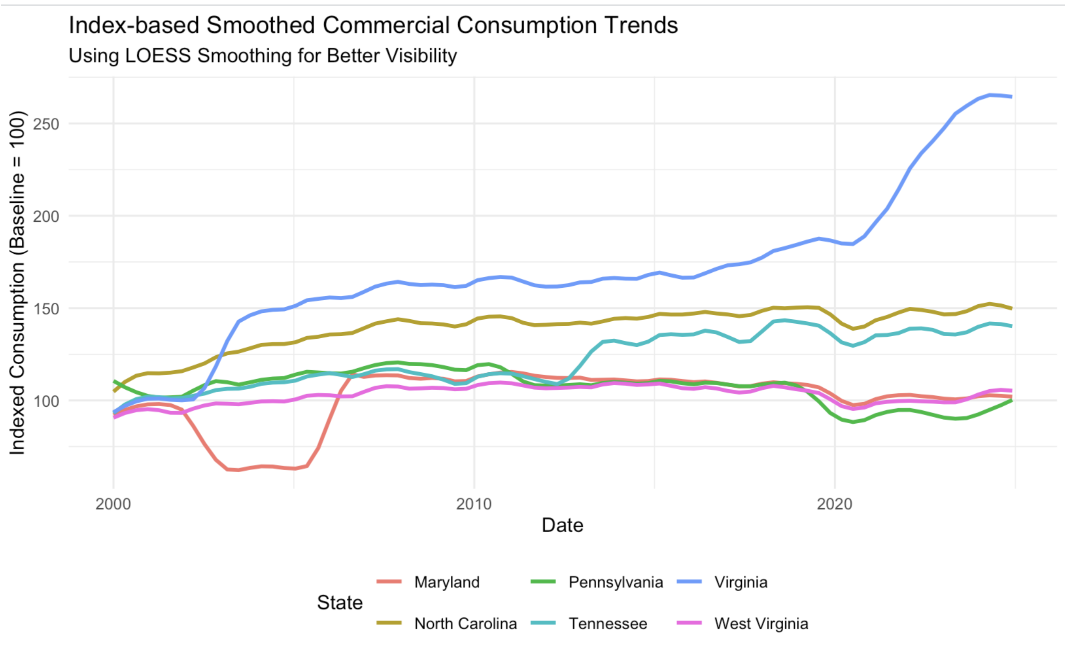


Figure 6: Index-based Smoothed Commercial Electricity Consumption across US Northeastern States. Data Source: (EIA, 2025a)

2 Literature Review

Despite their increasing importance, the academic literature linking data centres with power markets is limited. Since the 2000s, there have been limited and infrequent industry reports and academic papers. The existing literature primarily focuses on estimating data centre's electricity consumption and industry forecasts (Shehabi et al., 2016), (EPRI, 2024). (Shehabi et al., 2024). (Mytton and Ashtine, 2022), (Shehabi et al., 2018) and (Masanet, 2020) offer empirical evidence that energy efficiency improvements play the most important role in the long-term electricity consumption of a data centre. For example, Masanet (2020) estimates that power usage per unit of compute has annually decreased by 20% since 2010 due to new cooling technologies or hyperscaling. Such efficiency gains offset the overall nominal growth of the industry, resulting in a stable end-consumption across 2000-2016 (Shehabi et al., 2018). The trend of offsetting new demand through efficient computing is claimed to be broken by the broad popularisation of 'AI' tools, which are substantially more compute heavy. Several papers aim to isolate the impact of 'AI' services on electricity consumption of data centres. Verdecchia et al. (2023) and de Vries (2023) offer a critical review of the academic literature of 'AI'-based power demand, which demonstrates contemporary Large Language Models (LLMs) are orders-of-magnitude more power intensive than e.g. traditional search engines (Fig 7). EPRI (2024) and Koot and Wijnhoven (2021) highlight other power-intensive uses of computing infrastructure. These include cloud computing and data storage: heavily popularized during the Covid-19 pandemic, and cryptocurrencies. When combined, the rapid growth of these technologies is suggested to have caused major shifts in demand for computing infrastructure since at least 2020, leading to substantial electricity usage (Fig 24). The study seeks to verify such claims. However, the aforementioned literature focuses primarily on the estimation of the past and future data centre's electricity consumption, without considering its broader economic impacts[1]. The study seeks to fill this gap.

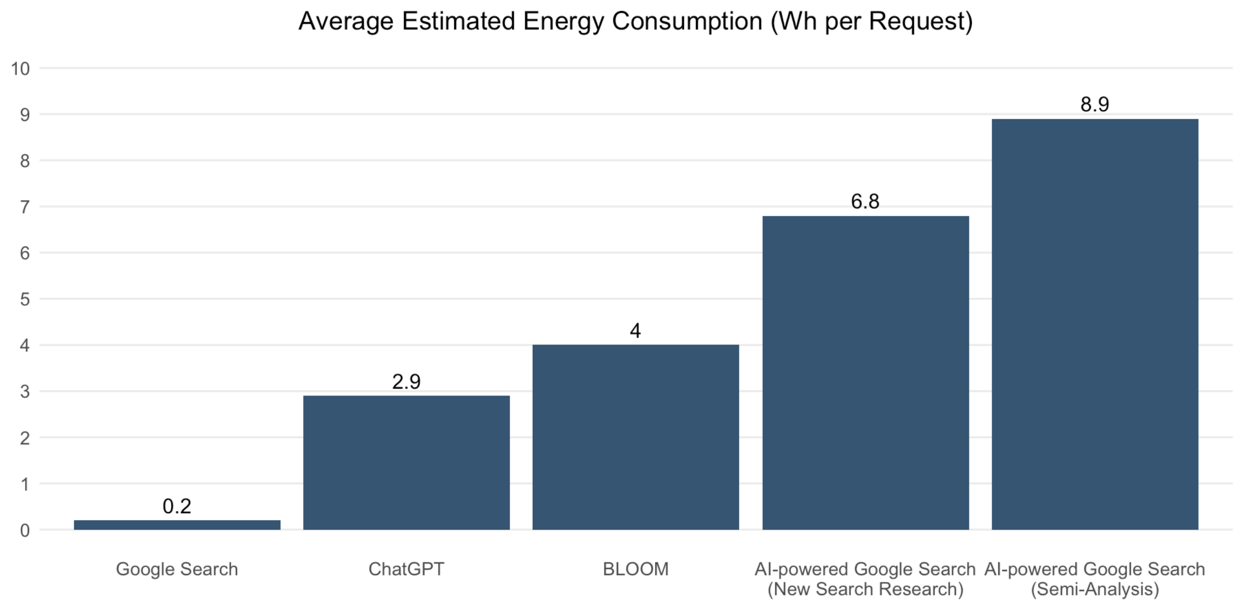


Figure 7: Average Power Consumption of various search models. Data Source: (IEA, 2024)

.

[1]For example: Shehabi et al. (2024) write the *"study seeks only to estimate the direct electricity use by data centres, not any underlying economic factors or transitions that may substantially change the underlying environment and technological base"*.

2.1 Existing Methodologies

This research draws from several existing literature, that employ the studies' chosen methodologies in the broader economic analysis of electricity markets. Borenstein et al. (2002) demonstrates that increased demand for electricity with constrained supply leads, to increased prices, an underlying mechanism investigated by the study. Li et al. (2021) directly employ structural break tests (Chow, 1960), (Andrews, 2003) to estimate the impact of oil shocks on residential (household) electricity consumption, while Esmaeili and Rafei (2021) deploy VAR analysis to deconstruct impact of the same oil shocks on electricity prices and overall inflation. Haldrup et al. (2010) investigate price convergence between different geographical power systems in the Nordics: electricity prices tend to spill over across highly integrated power markets. In the US context, (Stock and Watson, 1996) demonstrate the importance of incorporating structural break tests for long term times series analysis. Trost (1995) uses the restricted VAR to forecast regional monthly electricity prices and consumption providing useful insights into optimal choice of variables, while Horowitz (2007) presents empirical and theoretical evidence for spillover effect across state-level electricity markets. Similarly, Ros (2017) demonstrates the different elasticity levels across different sectors: residential consumers tend to be more responsive to price shocks than commercial and industrial power offtakers. Synthetic Control Analysis has seen a limited use, primarily to estimate the effects of restructuring of the US power markets after 1990s (Hill, 2021). This research seeks to leverage the aforementioned methodologies to investigate the extent to which data centres influence prices and demand across regions and non-commercial sectors: primarily residential and industrial[2].

2.2 The Case Study of Virginia

The majority of the existing literature focuses on either global or national levels only (Shehabi et al., 2024), (Mytton and Ashtine, 2022), (EPRI, 2024). However, data centres, due to investment requirements for close proximity to fibre networks, are clustered within specific geographic locations and should be analysed on regional level. Therefore, this research is grounded in the unique context of the state of Virginia. Turner (2025) and JLARC (2024) offer an official review of the development of data centre industry in Northern Virginia, particularly Loudoun County[3], and provide a review of existing policy incentives dedicated to data centres. Virginia, both on county and state levels, incorporates similar tax and permit policies to competing states, which reiterates its unique, exogenous, geographic feature Turner (2024), DoE (2024). Additionally, commercial literature offers two important insights for policymakers:

1. **Market saturation**: Virginia's power grid is at its full capacity, meaning additional investments in data centres are limited due to constrained power supply,

2. **Increased electricity prices**: If new power generation capacities are not built, substantial electricity price increases are forecast across all sectors in Virginia (Aurora Energy Research, 2024), (Dominion Energy, 2025).

The literature review reveals there exists a sufficient knowledge gap to study the impact of data centres on electricity markets, grounded in the regional context of Virginia US. The study seeks to fill this gap by deploying verified econometric methods to offer robust quantitative evidence of the impact of data centres on power markets within the region of Northeastern US and expand the economic dimension of contemporary literature on data centres. The chosen methodologies: structural break, synthetic control analysis and VAR, have been often deployed in the context of electricity markets motivating the choice of tools for this study.

[2] The Transportation sector is a minor power user and is disregarded in the study EIA (2025a).

[3] In 2008 Loudoun has introduced the first incentives to attract the industry, expanded in the late 2010s (Fig 17).

3 Methodology

3.1 Data

According to the author's knowledge, there exist no regular, official and reliable monthly or annual estimations of direct state-level electricity usage from data centres (Shehabi et al., 2024). Therefore, to estimate Virginia's data centre electricity consumption, this study constructs a new dataset based on several publicly available databases consisting of: macroeconomic indicators, weather and seasonality controls, granular electricity market data, and state-specific controls. The data range spans 2000-2024 at monthly frequency. It covers a regional cluster of neighbouring states that belong to the common regional electricity transmission system (the PJM): Maryland, North Carolina, Pennsylvania, Tennessee, Virginia and West Virginia[4]. Neither of the states, apart from Virginia, have any major capacity of computing infrastructure (EIA, 2024a), which differentiates them from Virginia. In 2003, the EIA has changed its electricity consumption accounting methods (EIA, 2024b), which changed the data structures primarily for Virginia and Maryland. Additionally, the US power markets have experienced a series of restructurings until early 2005. Therefore, the final analysis covers the years 2008-2024 which offer greater number of available variables, coherent data accounting and does not incorporate long term effects of market restructuring (DoJ, 2008). The electricity data is available at sector levels: residential, commercial, industrial and transportation, while macroeconomic indicators can be deconstructed to specific industries, e.g. employment in information sector. Crucially, data centres are classified within the 'commercial sector' category of electricity usage (EIA, 2024a). This allows to isolate their consumption from other power-demanding industries, such as manufacturing, and from impact of recent US industrial policies e.g., the IRA which are primarily categorised in the 'industrial sector' (Bistline et al., 2023).

3.2 Identifying Structural Breaks

I aim to reproduce an estimate of Virginia's power consumption of data centres, which is based on the assumption that data centres are primarily located based on exogenous geographic proximity to internet fibre networks (EPRI, 2024), (Turner, 2024). I first investigate whether there was a statistically significant structural break in Virginia's commercial power consumption, i.e. whether there was a sufficiently strong change in the electricity demand pattern. The standard approach to structural breaks is introduced by Bai and Perron (1998), covering an example linear model:

$$y_t = x_t'\beta + z_t'\delta_j + \epsilon_t \text{ where:}$$

1. y_t is the independent variable,
2. $x_t(p \times 1)$ and $z_t(q \times 1)$ are vectors of covariates,
3. β which is the corresponding vector of constant coefficients,
4. δ_j which is the corresponding vector of regime-dependent coefficients,
5. ϵ_t is the error term,
6. $(T_1, \ldots, T_m)$ are the structural breaks,
7. $(j = 1, \ldots, m + 1)$ are the existing regimes in the total timespan: $(T_0, \ldots, T)$

[4] Additionally, I incorporate two coastline states: Georgia and South Carolina, for control purposes.

Table 1: Variables Table for Model specifications – each variable is available for each state i[5].

Variable	**Name**	**Type**	**Source**
Time Variables			
date	date	Date	
Year	Year	Numeric	
Month	Month	Numeric	
State-level Electricity Variables (Monthly)			
i_commercial_consumption	Commercial Electricity Consumption	GWh	(EIA, 2025a)
i_commercial_price	Virginia Commercial Price	Cents/kWh	(EIA, 2025a)
i_nat_gas_price	Natural Gas Price	USD per Thousand Cubic Feet	(EIA, 2025c)
generation_comm_small_solar	Small Scale Solar PV Generation	GWh	(EIA, 2025)
i_cdd	Cooling Degree Days	Numeric	(NOAA, 2025)
i_hdd	Heating Degree Days	Numeric	(NOAA, 2025)
i_avg_temp	Average Monthly Temperature	Fahrenheit	(NOAA, 2025)
Economic Indicators (Monthly)			
interest_rates	Federal Funds Effective Rate	Percent	
i_coincided_index	State Coincided Indexes - a GDP Proxy	Index	(FED Philadelphia, 2025)
i_employ_month	Monthly Non-Farm Employment	in thousands	(FRED St Louis, 2025c)
i_information_employ	Information Sector Nominal Employment	in thousands	(FRED St Louis, 2025c)
i_MFGHRS	Average weekly hours, private employees	Nominal	(FRED St Louis, 2025c)
i_non_manf_exp	Exports of Non-Manufactured Commodities	Millions of Dollars	(FRED St Louis, 2025b)
i_imports	Imports of Goods	Millions of Dollars	(FRED St Louis, 2025b)
i_business_app	Business Applications with Planned Wages:	Nominal	(FRED St Louis, 2025a)
VA_bvix_index	Average Price of Bloomberg Virginia (BVIX) market index	USD-denominated market index	(Bloomberg L.P., 2025)
Dummy Variables			
Seasonal Monthly Dummies	Dummy per each month: January–November	Bivariate	
Covid Lockdown	March–May 2020	Bivariate	
Covid Reopen	June 2020 – October 2022	Bivariate	
Covid Second Wave	November 2020 – March 2021	Bivariate	
Covid Recovery	April 2021 – April 2022	Bivariate	
Global Energy Crisis	March 2022 – December 2022	Bivariate	

Exceptions include national-level variables, such as interest rates, and state-dedicated variables, such as BVIX stock index.

Following this framework, I introduce three methods to test whether Virginia's commercial electricity consumption has experienced a structural break between 2000-2024.

The Andrews Test (Andrews, 2003) is the method of choice based on its two properties: it is designed for *testing for structural instability over a short time interval, particularly at the end of a data sample* (*Ibidem*) and for identifying structural breaks when the timing of the break is uncertain (Fair, 2003). The Andrews Test searches for the most likely structural break point by computing the S-test statistic - a variant of the standard F-statistic (Chow, 1960), (Andrews, 2003), over a range of potential breakpoint windows[6], to identify the most probable date. The study's regression can be represented as:

$$
\begin{aligned}
y_t = {} & \beta_0 + \beta_1 \text{commercial prices} + \beta_2 \text{natural gas prices} + \beta_i \text{weather controls} \\
& + \beta_i \text{macroeconomic controls } + \beta_i \text{COVID-19 Dummies} + \beta_i \text{Energy Crisis} + \\
& \gamma_i \text{Seasonal Dummies} + \sum_i \delta_i \text{lag variables}_i + \epsilon_t
\end{aligned}
$$

Given the high seasonality and persistent temporal effects of the underlying data and variables, the study uses the Autoregressive Distributed Lag (ARDL) to determine the optimal lag structure, defined as a multivariate ARDL model (Fig 6.2.):

$$
\begin{gathered}
y_t = \alpha_0 + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=0}^{q} \beta_j X_{t-j} + \epsilon_t \quad \text{such that:} \\
\min_{\{p,q\}} \text{AIC}(p, q) = \ln(\hat{\sigma}^2) + \tfrac{2k}{n} \quad \text{where:}
\end{gathered}
$$

1. p and q are optimal lags for the dependent (y_t) and independent variables (Vector X_t), $\hat{\sigma}^2$ is the estimated variance of residuals,
2. k is the number of estimated parameters,
3. n is the number of observations.

The model uses the Akaike Information Criterion (AIC) (Akaike, 1974), which is more appropriate than e.g. Bayesian Information Criterion (BIC), as it is a flexible method to accommodate longer lag structures in a smaller dataset. The Andrews Test fits with the study's timing - the impact of data centres on electricity markets has only been recent, and removes bias towards specific pre-defined dates (Andrews, 2003). The method is verified by the Chow Test (Chow, 1960), which incorporates the pre-defined candidate-date structural break. It checks for a null hypothesis of whether there was no shift in coefficients before and after the chosen breakpoint.

$$
H_0 : \boldsymbol{\beta}_t = \boldsymbol{\beta} \quad \text{for all } t \quad \text{vs.} \quad H_1 : \boldsymbol{\beta}_t = \begin{cases} \boldsymbol{\beta}_1 & \text{for } t < T_{break} \\ \boldsymbol{\beta}_2 & \text{for } t \geq T_{break} \end{cases}
$$

The test is performed by estimating three regressions:

1. The unrestricted models for two subsamples (pre- and post-break),
2. The restricted model on the pooled data (assuming no break) where:

$$
\begin{gathered}
y_t = x_t' \beta_1 + \epsilon_t, \text{ for } t \leq T_{break}. \\
\text{and} \\
y_t = x_t' \beta_2 + \epsilon_t, \text{ for } t > T_{break}.
\end{gathered}
$$

[6] defined through parameter h Andrews (2003).

where the T_{break} is a pre-defined break date.

A f-statistic is then computed to evaluate the statistical significance of the breakpoint (see Appendix):

$$F = \frac{(SSR_r - (SSR_1 + SSR_2))/k}{(SSR_1 + SSR_2)/(n_1 + n_2 - 2k)}$$

Finally, for further robustness, I perform the Cumulative Sum Control (CUSUM) (Brown et al., 1975), to test for the coefficients stability over time and indicate whether the shock was abrupt or smooth. The CUSUM test is based on the intuition that if a regression's β_i change over time, then a one-period-ahead forecast will not be accurate. If the regression's coefficients remain structurally unstable, the cumulative sum of recursive residuals w_i will increase instead of stabilising. Therefore, we can formally define a two-sided CUSUM as a sequential sum:
$CUSUM_t = \sum_{i=1}^{t} \hat{w}_i$.

3.3 Synthetic Control Analysis

I then proceed to the synthetic control analysis to demonstrate that data centres perform a crucial role in Virginia's structural break. First proposed by Abadie and Gardeazabal (2003), the method constructs a 'synthetic' control region, *which resembles the economic characteristics of the study region*. It then follows the economic evolution of such 'counterfactual' control index: *syntheticVirginia*, and is then compared with the actual region: *realVirginia*. The index region, which does not incorporate controls for data centres, is then compared to identify the impact of the explanatory variable, in this case: the computing industry, on the dependent variable: commercial electricity consumption. The research follows the guidelines set by (Abadie, 2021). I specifically choose the Synthetic Control Analysis method as it is often leveraged to study evolutions between different geographies (Abadie and Vives-i Bastida, 2021) and it simultaneously accounts for major macroeconomic effects such as the COVID-19 pandemic, which equally affected both Virginia and neighbouring states. The treatment effect is defined as:

$$Y_{1t} = Y_{1t}^{Synth} + \tau_{1t} D_{1t}$$

where:

1. Y_{it} is the outcome variable of interest for state i at time t, where $i = 1$ corresponds to the treated state (Virginia), and $i = 2, \ldots, J+1$ represent the *donor pool* (Control States) (Abadie, 2021), (Ponne, 2023). Therefore, Y_{1t}^{Synth} is the potential outcome for Virginia in the absence of treatment - no data centres,

2. τ_{1t} is the treatment effect

3. D_{1t} is a binary indicator equal to 1 if $t \geq T_0$ (post-treatment period), and 0 otherwise.

In comparison to standard difference-in-difference, the intuition behind Synthetic Control Analysis is that a combined set of states from the donor pool approximates the affected state (Virginia) much more efficiently. Therefore, the counterfactual Y_{1t}^{Synth} is estimated by a weighted average of the control units:

$$\hat{Y}_{1t}^{N} = \sum_{j=2}^{J+1} w_j Y_{jt}$$

The weights $W = (w_2, \ldots, w_{J+1})'$ are conditioned on $w_j \geq 0$ and $\sum_{j=2}^{J+1} w_j = 1$, and optimised (Abadie et al., 2010) such that:

$$\sum_{j=2}^{J+1} w_j^* X_j \approx X_1$$

with X_i being vector of pre-treatment characteristics for state $i = 1$. The difference between the treated and synthetic control is then computed as:

$$\hat{\tau}_{1t} = Y_{1t} - \hat{Y}_{1t}^{Synth}$$

Importantly, due to the methodology's design the impact of small shocks when the underlying variable is highly volatile - such as highly seasonal electricity consumption (Fig 22), is limited. Therefore, only strong and persistent shocks will be recognised by the model (Abadie, 2021). Therefore, the second null hypothesis is:
H_0 :*Virginia's commercial electricity demand will be higher than the synthetic region control after the structural break point.*

3.4 Vector Autoregression

Following the structural break and synthetic control analyses, I investigate the dynamic interdependencies between Virginia's commercial electricity consumption and the broader regional electricity market. I deploy the Vector Autoregressive Model (*VAR*) (Sims, 1980) to capture and account for bidirectional relationships between e.g., electricity consumption and electricity prices across industrial and commercial sectors, without requiring strong a priori causal assumptions. The reduced-form VAR(p) model for a vector of N endogenous variables $\mathbf{y}_t$ is defined as:

$$\mathbf{y}_t = \mathbf{A}1\mathbf{y}t - 1 + \mathbf{A}2\mathbf{y}_{t-2} + \cdots + \mathbf{A}p\mathbf{y}_{t-p} + \boldsymbol{\epsilon}_t, \quad \text{where} \quad \boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \Sigma), \text{ and:}$$

1. $\mathbf{y}_t$ is a $(N \times 1)$ vector of time-series variables at time t,
2. $\mathbf{A}_i$ are $(N \times N)$ coefficient matrices capturing the effect of lag i,
3. $\boldsymbol{\epsilon}_t$ is a N-dimensional iid white noise vector.

The model is then estimated using OLS per each equation. The specified models consist of 5 variables:

1. $X_{j,s,t}$: Electricity *price* or *consumption* in the target $state_j$ and $sector_s$ - the target variables,
2. $\hat{\tau}_{1t}$: Synthetic Residuals obtained through SCA, which represent the proxy for Virginia's data centres' electricity consumption,
3. $GDP_{j,s,t}$: GDP in the target state - a proxy for economic activity (Kraft and Kraft, 1978),
4. $Temp_{j,s,t}$: Average Monthly Temperature in the target state - a proxy for seasonality,
5. $Gas_{j,s,t}$: Natural Gas Price in the target state - natural gas price is the marginal electricity price setter in the US (EIA, 2025b).

To determine the optimal lag length p and to ensure cohesive methodology across the study, the $Var(p)$ model uses the Akaike Information Criterion (AIC) for the same reasons as the ARDL.

To ensure robustnesses of the tests, I use the augmented Dickey–Fuller (ADF) to test for stationarity of the variables. The Dickey-Full test results suggest all control variables are stationary apart from the main variable of interest: *synthetic residuals* - a proxy for Virginia's data centres electricity consumption.

Table 2: ADF Test Results for Stationarity

Variable	Statistic	P_Value	Lag
VA_commercial_price	-4.22	0.01	3
VA_coincided_index	-3.62	0.03	3
VA_nat_gas_price	-3.85	0.018	3
VA_avg_temp	-14.51	0.01	3
synth_residuals	-0.81	0.96	3

Therefore, since non-stationarity violates the initial VAR assumptions and leads to spurious regressions (Granger and Newbold, 1974) I use first-differences:

$$\Delta Y_t = Y_t - Y_{t-1}$$

to model the VAR. The Johansen test (see table 5) ensures no co-integration which excludes the need to perform a Vector Error Correction Model. To interpret, the dynamic behaviour of the system, I deploy Impulse Response Functions (IRFs) which model the effect of a one-unit shock on the current and future values of $\mathbf{y}_t$, while accounting for lagged effects of e.g., economic activity. The IRF is particularly useful to assess the magnitude of the impact of the variable of interest, while incorporating exogenous influences of e.g. seasonal weather patterns and macroeconomic performance. Therefore, the final hypotheses are:

1. H_o A positive shock in synthetic Residuals ($\hat{\tau}_{1t}$) will increase electricity prices across Virginia's and regional power sectors.

2. H_o A positive shock in synthetic Residuals ($\hat{\tau}_{1t}$) will decrease electricity consumption across Virginia's and regional power sectors.

4 Results

I deploy several configurations of the Andrews Structural Break Test (Andrews, 2003), which yield statistically significant results. I then verify the suggested breakpoint dates with Chow Test (Chow, 1960). The resulting summary table is:

Table 3: Andrews and Chow Tests for structural breaks in Virginia's Commercial Electricity consumption

Model	R-squared	Break Date(s)	Break Test p-value
Andrews Test (2008–2024)	0.95	2021-05	<0.01***
Chow Test (2008–2024)	0.96	2021-05	<0.01***

The tests suggest there was a structural upward break in Virginia's commercial electricity consumption in May 2021, which supports the stated hypothesis and previous literature (EPRI, 2024). The residuals of each model - which deliberately does not account for data centres - tend to increase after 2020 (Fig 25). Similarly, the CUSUM test indicates a stable growth of the model's error rates, until achieving statistically significant instability of the coefficients after 2021[7]. This demonstrates a growing omitted variable bias when modelling Virginia's commercial electricity consumption without incorporating a dedicated proxy for e.g., data centres.

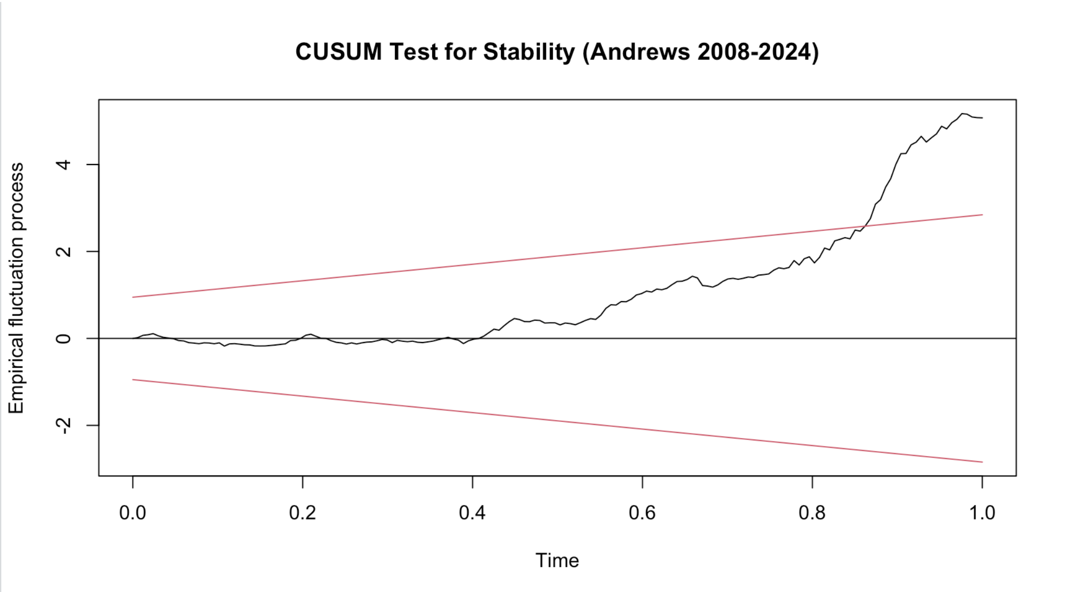


Figure 8: The CUSUM test with the 95% confidence intervals

The Synthetic Control Analysis (SCA) (Fig. 9) leverages Virginia's geography: unique close access to international fibre optic cables, as a quasi-exogenous property for data centre buildout within the Northern Virginia region. Such a framework assumes the only main difference between Virginia and surrounding states is the presence of data centres. Therefore, the difference between *Real* and *Synthetic* commercial electricity consumption can be assigned to data centres combined with a structural break where:

$$Synthetic\ Gap = RealVirginia_t - SyntheticVirginia_t$$

[7]the 85th percentile of the sample set

The results follow the historical dynamics described by the literature: during 2000-2015 digitalisation[8] was offset by efficiency improvements resulting in stable consumption[9] (Shehabi et al., 2018). Since 2016, the data centres infrastructure has slowly started to grow, until the major shocks of Covid-19. The massive and rapid increase in demand for online data storage and internet traffic transferred to electricity consumption (Turner, 2024). Finally, the rapid popularisation of compute heavy LLM-tools such as OpenAI (2025) or Anthropic (2025) has maintained and extended the Covid-levels of electricity consumption (Shehabi et al., 2024). The Synthetic Control Analysis aims to capture such quasi-exogenous shocks: buildout, Covid-19 and LLM popularisation, and estimate the 'no-data-centres' scenario through the indexed non-Virginia states.

Crucially, the advanced and institutional review by EPRI (2024) offers a reliable and technical estimate of Virginia's data centre electricity consumption for the year 2023 at 33 TWh. The synthetic gap when aggregated to annual data is equal to 28 TWh for 2023, only a 15% mismatch. This yields two important results. First, the experiment tends to predict market dynamics in line with official estimates (*Ibidem.)* and existing literature (Fig. 23 via Shehabi et al. (2024)). Second, the *Synthetic Gap* underestimates rather than overestimates the data centre's consumption. This limits the risk of omitted variable bias. For example, some unobserved effects e.g., increased investments from the data centre industry, could lead to increased commercial activity and electricity consumption, which would be incorporated into the gap, resulting in an overestimated proxy for data centres power consumption. Verifying the proposed proxy with (EPRI, 2024) increases the robustness of the study's findings. I run an additional SCA for Tennessee (Fig. 25) which yields no major difference between *Synthetic* and *Real* Tennessee. Such Placebo Test, increases the robustness of this study's methodology and supports accepting the second null hypothesis.

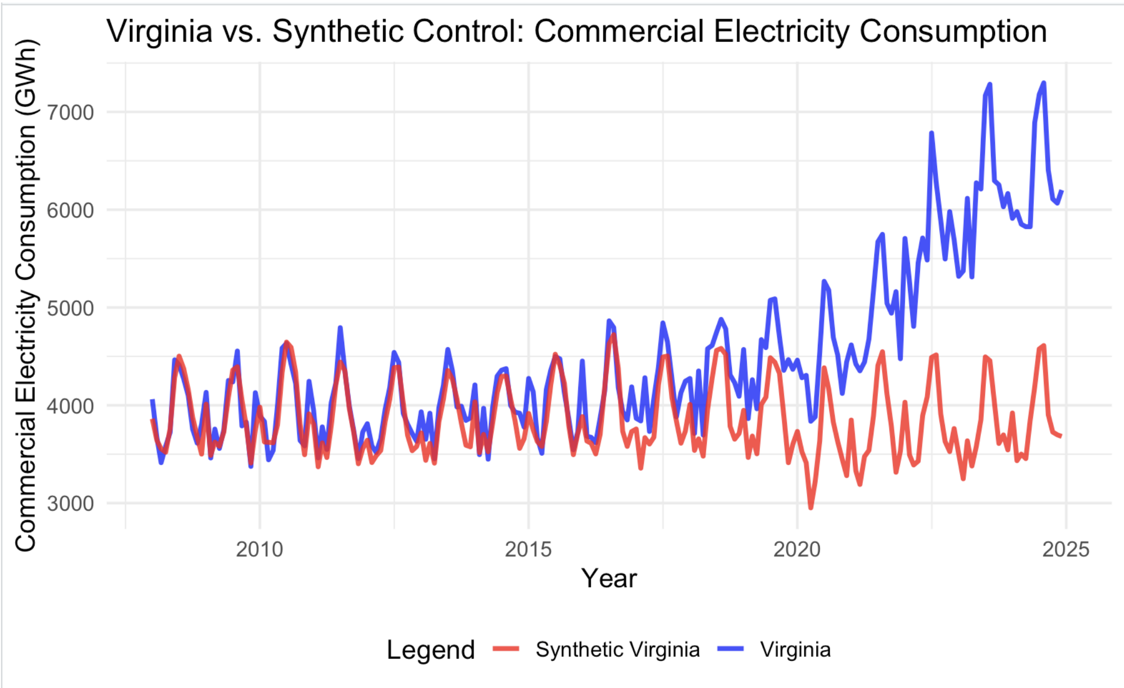


Figure 9: The Synthetic Control Analysis recreates the pathway of *targeted state* based on weighted features of *donor pool* states.

[8]I.e. growing internet usage, which was served by data centres. For example, Virginia's computing industry handled 70% of global internet traffic in 2019 (Virginia Economic Review, 2019).

[9]For example, Shehabi et al. (2018) write: *The trend in data centre electricity use since 2000 is a success story of energy efficiency.*

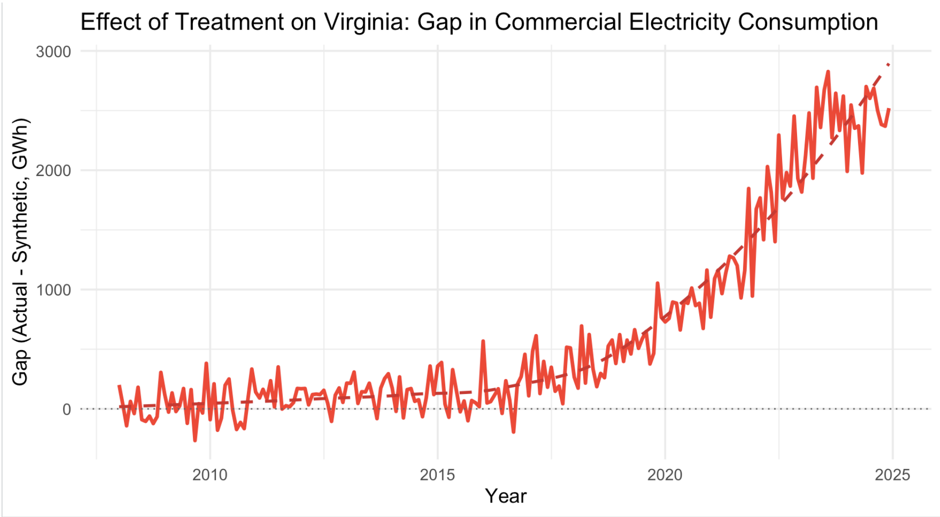


Figure 10: The gap between *Real* and *Synthetic* Virginia has been rising since at least 2016.

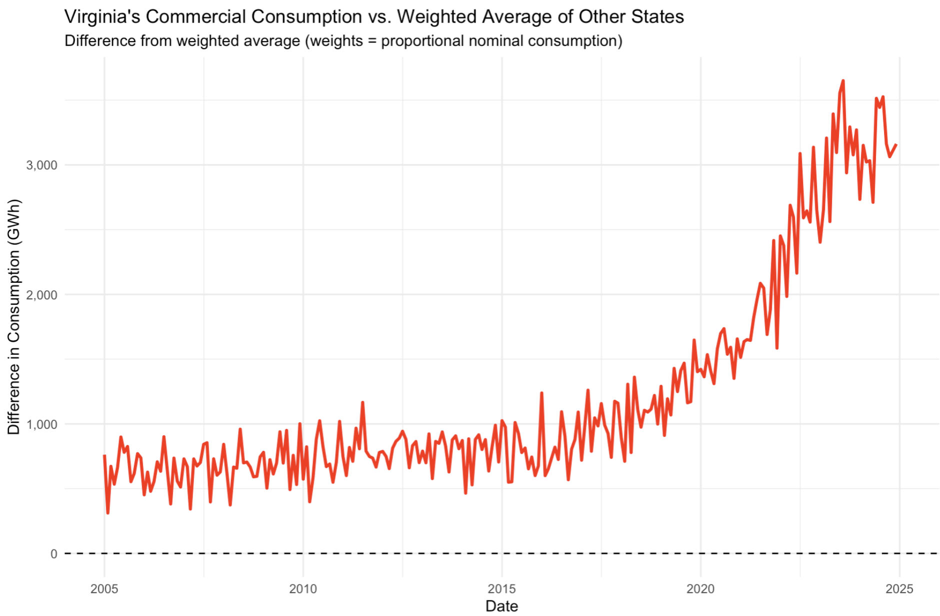


Figure 11: The energy usage gap between *Synthetic* and *Real* is in line with the mean-averaged comparison between Virginia and neighbouring PJM States.

I deploy the VAR model with 'Synthetic Residuals' based on the synthetic gap. Therefore, I only estimate the direct impact of approximated electricity consumption of data centres onto sectoral and regional electricity markets. I operate in GWh of consumed electricity, dollars per GWh prices and logged prices. Therefore, the Impulse Response Functions present an estimated impact of an increase in 1 unit of power consumption (1 GWh) on electricity price (Dollars/GWh and % percentage points) within a 24 months window. The intuitive theoretical mechanism is based on Haldrup et al. (2010): increased demand for electricity combined with constrained supply leads to higher prices. The analysis of Virginia's electricity markets, reveals statistically significant evidence for data centres' power consumption influencing sectoral electricity prices (Fig. 12). The IRFs demonstrate a sectoral differentiation: the impact on the industrial sector is more profound and persistent than for the residential sector. This is in line with Ros (2017), which argues residential electricity usage tends to be more seasonal - which increases the difficulty to isolate the shock effects of the IRF, and more responsive (elastic) to price shocks. Furthermore, the IRF for the industrial sector might be more substantial than for others due to the baseline effect - residential and commercial sectors consume much more power therefore, a nominal impact of 1GWh will be smaller.

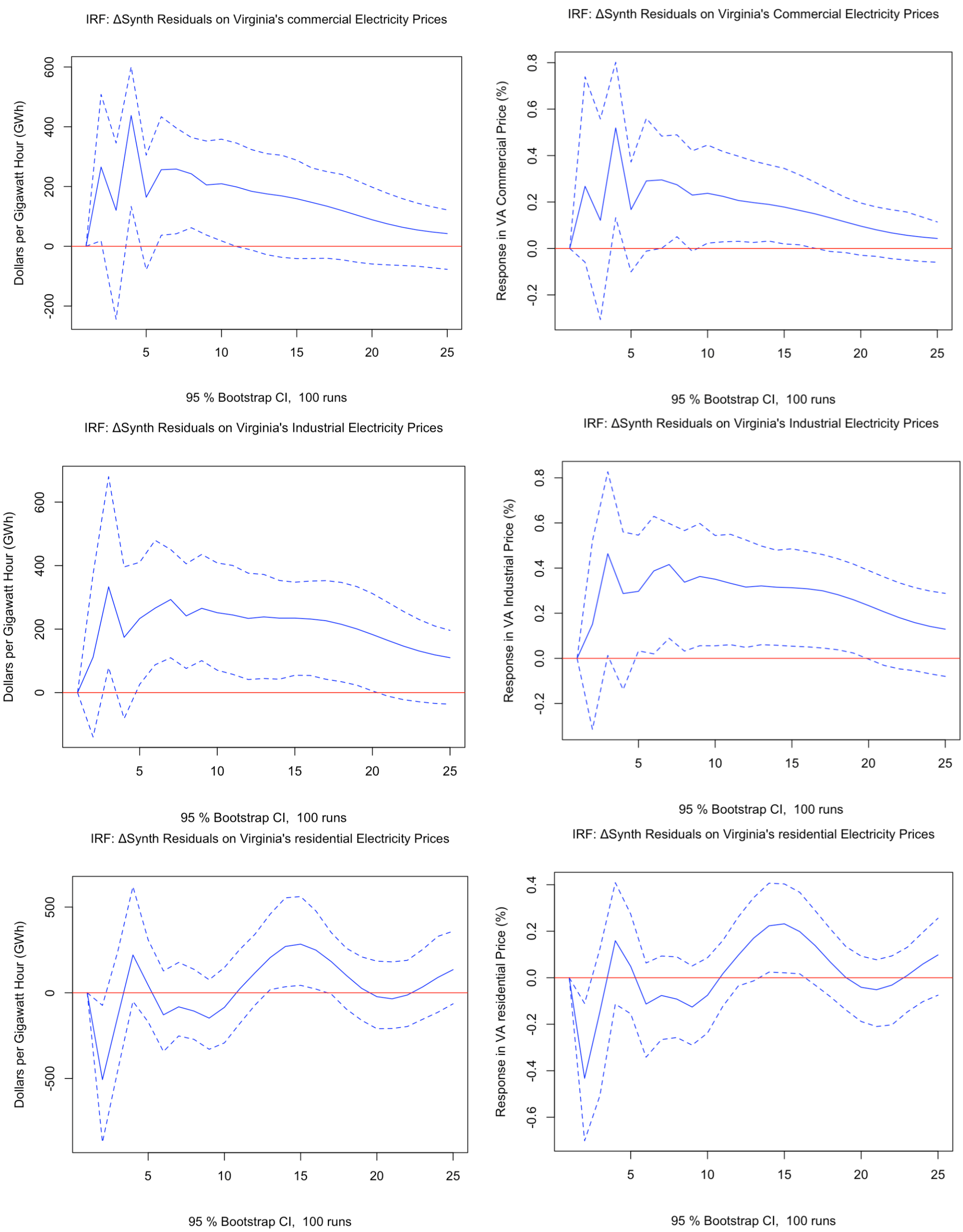


Figure 12: An increase of 1 GWh from data centres consumption influences prices across Virginia's electricity markets, with diminishing spillover effects

Given that the monthly 1GWh consumption corresponds to only 1.4MW of installed capacity[10], a, on average, 0.2% price increase is substantial. This demonstrates the scale of market saturation: a minor increase in additional demand can lead to major price fluctuations. JLARC (2024) predicts currently there exist 1500 MW of data centres under construction in Virginia. Based on the initial IRF estimates, such a sustained demand shock could lead to a compounded price increase of over 20% in the most affected sectors, primarily industrial, if not met with a proportional increase in electricity supply.

There are two states which are net electricity exporters and operate within the same electro-energetic system (the PJM) as Virginia: Pennsylvania and West Virginia. I investigate whether the energy flows between the states and Virginia (EIA, 2024) have resulted in some effects on their domestic markets. The experiment suggests data centres, in the form of *synthetic residuals* proxy, could result in spillover effects influencing the prices in neighbouring states of Pennsylvania and West Virginia[11]:

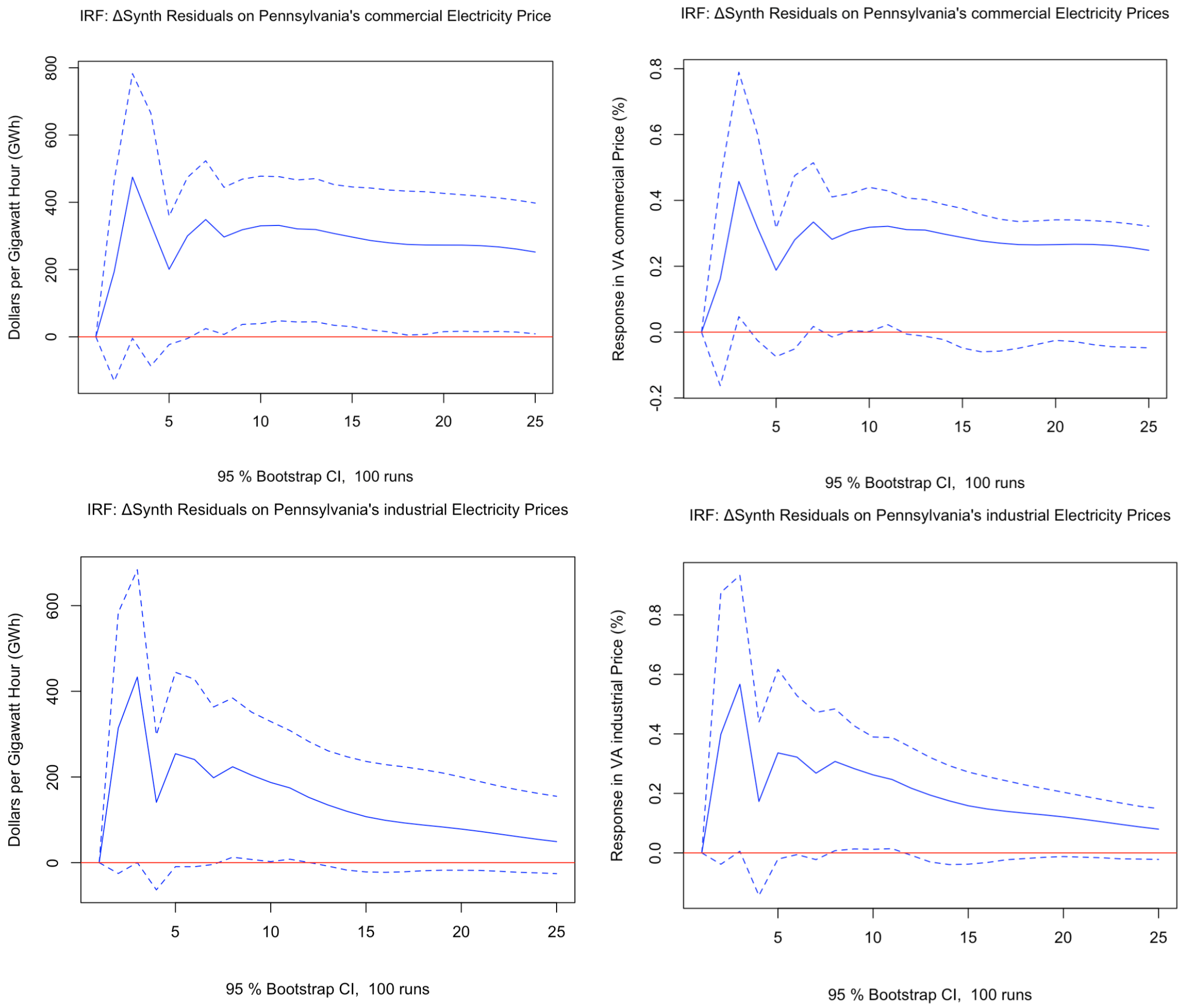


Figure 13: There is a statistically significant impact of increased electricity consumption from data centres onto Pennsylvania's sectoral electricity prices.

[10] 1.4MW is a minimum value assuming 100% load factor.

[11] Residential sector is the least empirically and theoretically responsive to IRF shocks, therefore I omit it in the interstate analysis

The impact of data centres on West Virginia's electricity markets is less visible for one potential reasons. The state's main energy source is coal (85% of energy mix) EIA (2024b), meaning that coal price fluctuations substantially influence electricity prices. The VAR model is fitted for natural gas prices, leading to a potential ill-fit[12]. This limits the potential to isolate the true impact onto West Virginia's electricity markets.

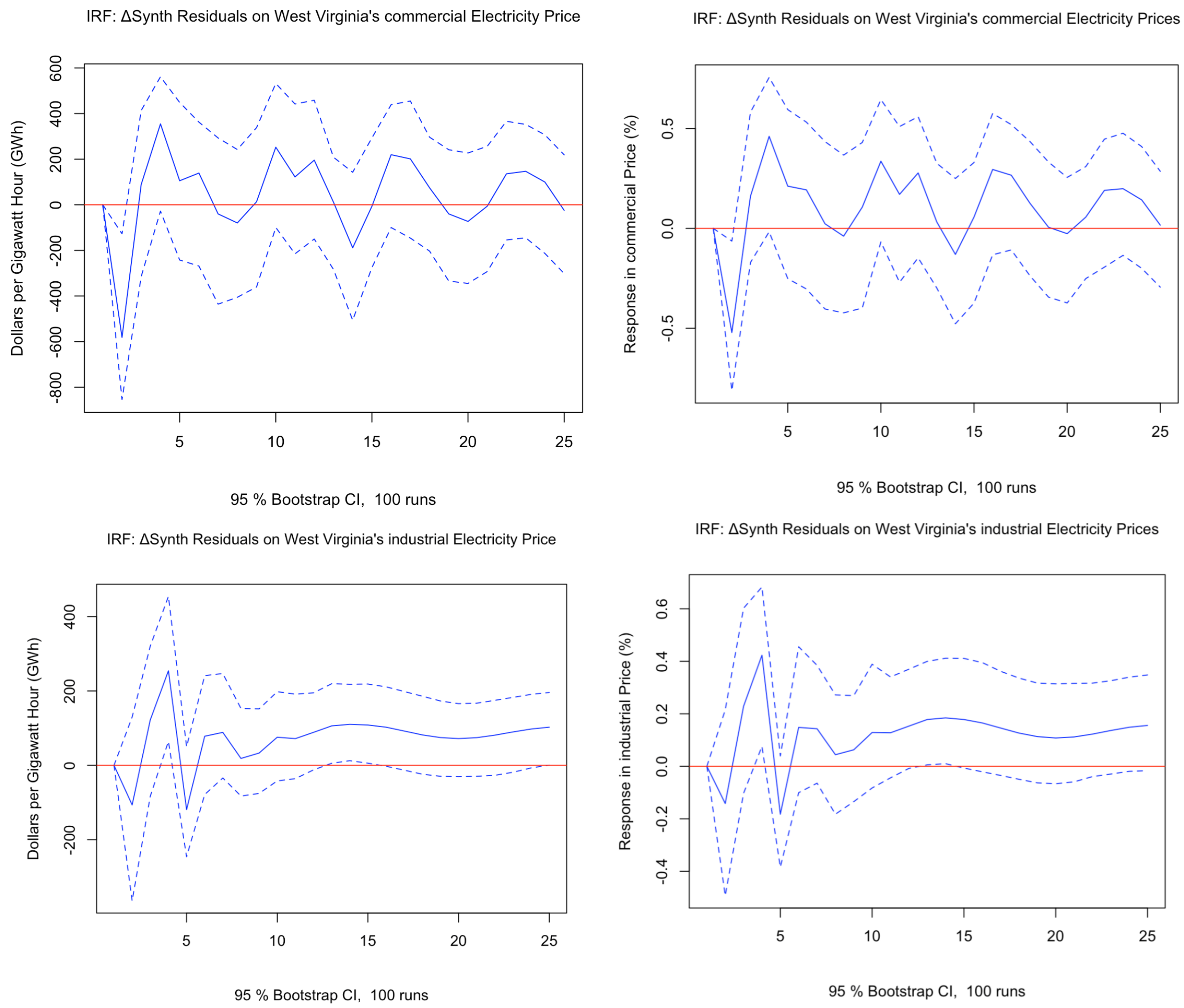


Figure 14: There is no statistically significant impact of increased commercial consumption from data centres onto West Virginia's sectoral electricity prices.

For control purposes, I then test the hypothetical impact of Virginia's data centres on commercial electricity markets of South Carolina and Georgia - two states in close proximity to Virginia, yet not integrated with its electrical grid (Fig. 15). Since both states are not members of the PJM, the hypothetical impact should not be expected. The exercise yields no statistical significance, increasing the robustness of chosen methodologies, and suggesting the hypothetical IRF shocks may actually be driven by data centres and not other unobserved effects.

[12]Unfortunately, EIA does not offer direct, historical coal prices for West Virginia.

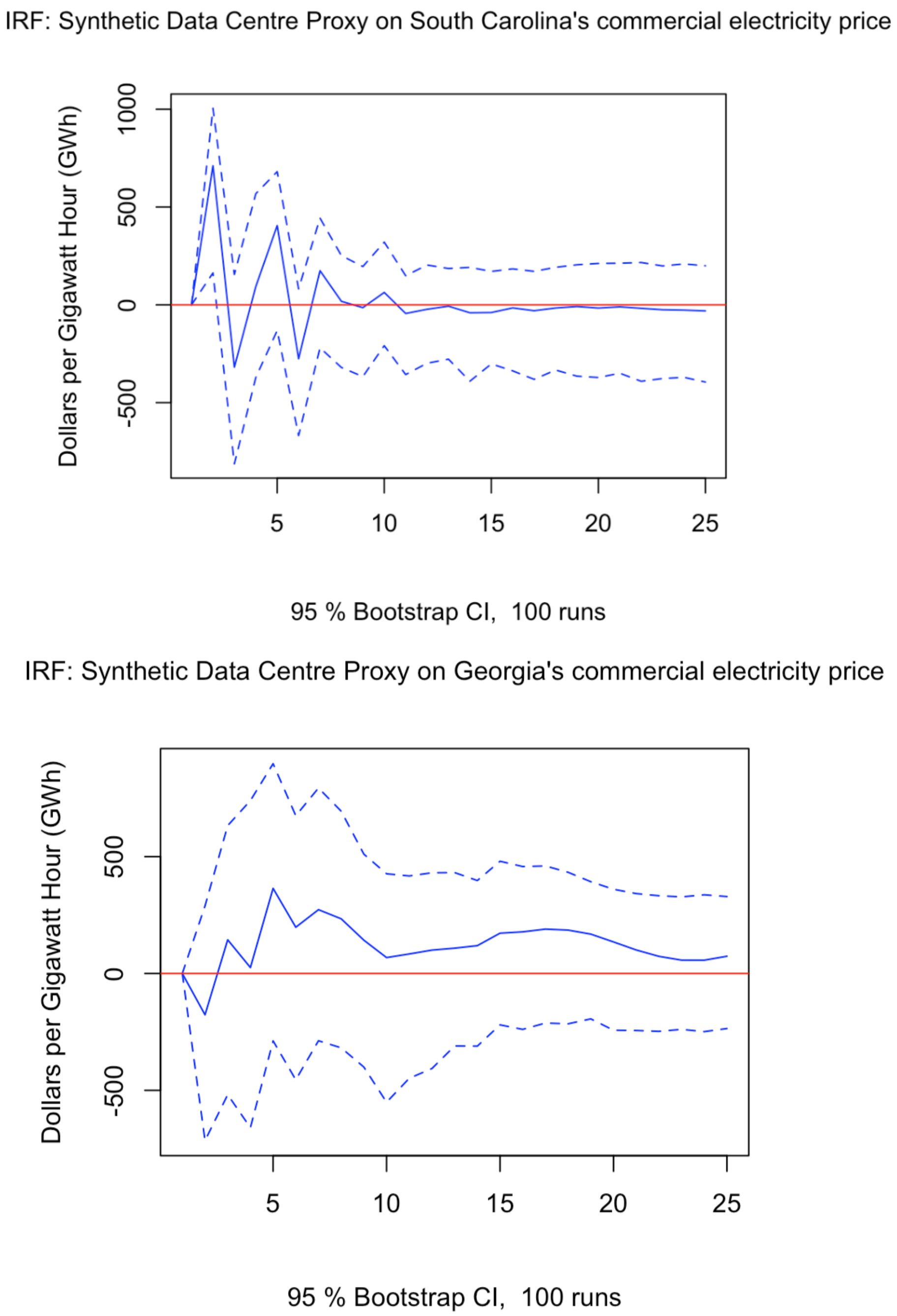


Figure 15: The hypothetical impact of Virginia's data centres on disconnected markets: South Carolina and Georgia, is non-significant, ensuring the model is correctly deployed.

The experiment provides no evidence for data centres power usage influencing electricity consumption patterns across other sectors and states. The IRF reveals high seasonality of consumption patterns, particularly of the residential sector, which may limit the possibility to isolate the shock to demand. Furthermore, the limited impact may be due to the baseline effect: Virginia's average monthly commercial consumption is equal to 6300 GWh, therefore a 1 GWh shock may be too small to any strong effects. Therefore, I reject the last, null hypothesis. Due to the limitations of this work, additional theory and empirical experiments are required to further verify the impact of data centres onto consumption patterns.

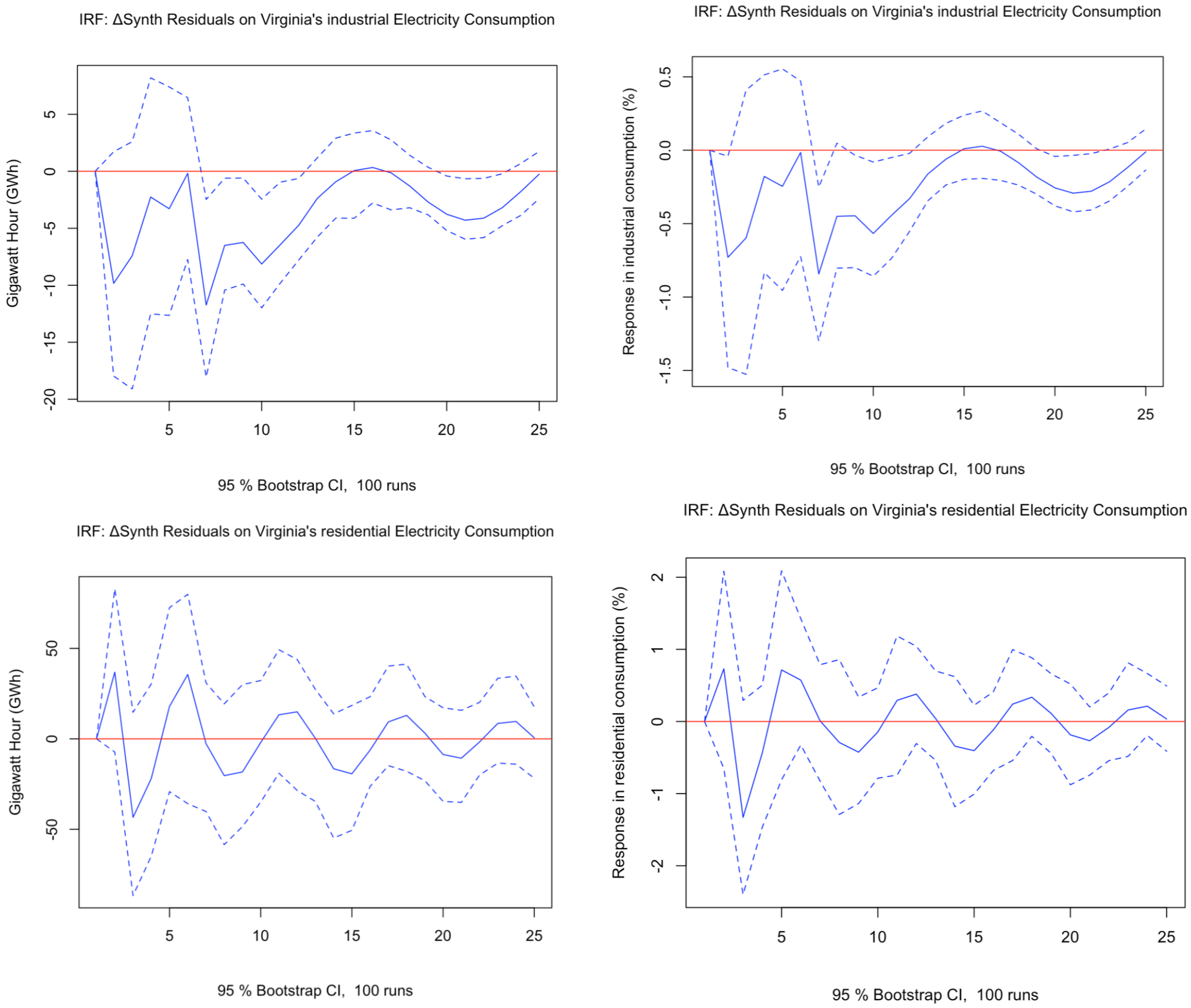


Figure 16: The proxy for data centres power usage does not create statistically significant impact onto Virginia's electricity consumption.

5 Conclusion and Policy Implications

Technological and economic benefits of the computing industry are well-studied and promoted on national and global levels (EPRI, 2024). This study offers a regional analysis of some of the downsides and costs they generate: increased electricity prices which may lead to market saturation.

5.1 Policy Implications

The study offers an econometric method to estimate electricity consumption of data centres in the state of Virginia. It estimates the impact of the industry on broader power markets across the regional states via a dedicated proxy. A positive exogenous shock of 1 GWh of data centres electricity consumption corresponds on average to a 0.2% persistent, commercial price increase in Virginia. This indicates a major market saturation and is supported by further evidence of potential spillover effects across regional states and sectors. In a scenario of limited supply of additional power, the current growth of Virginia's computing infrastructure could substantially price-pressure non-data-centre industries and residential consumers (JLARC, 2024). This presents a major policy challenge for both Virginia's and regional lawmakers for several reasons:

1. **Demand Forecasting:** accurate predictions of future electricity demand are extremely important to grid operators, ensuring a stable and reliable power grid. Structural breaks play a major role in forecast adjustment (Stock and Watson, 1996). By identifying specific structural breaks and the interactions between the data centre industry and electricity markets, the study offers potential improvements to existing forecasting methodologies.

2. **Cost Benefit Analysis**: While the computing industry is responsible for more than half of external investments into Virginia (Turner, 2025), rising prices of electricity will lead to increased costs of living and perhaps even broader inflationary pressure (Esmaeili and Rafei, 2021). This research provides one of the necessary estimates to perform a broad cost-benefit analysis of Virginia's data centre industry.

   (a) **Environmental footprint**: the study's proxy offers first reliable estimation of monthly electricity usage of data centres, allowing for early, regional, estimation of their environmental footprint (Malmodin et al., 2024): a key component of contemporary cost-benefit analysis. Additionally, while this study incorporates only the usage of electricity, data centres are major consumer of other resources: primarily water (de Vries, 2023), which should be incorporated in future studies.

   (b) **Costs of infrastructure**: Expanding and sustaining Virginia's data centre industry will require substantial, regional, investments in power transmission infrastructure. This opens a major coordination problem: which federal, state or industry actors should finance the expansion of the grid? This research offers additional tools and estimates for policymakers to make informed decision on optimal costs allocations for the power grid expansion.

3. **Price protections**: The research provides some empirical evidence, power usage of Virginia's data centres spills over to neighbouring states. State-level policymakers in e.g., Pennsylvania, should analyse to what extent they ought to accept (or regulate) such increased electricity prices, which can undermine their own industrial and commercial sectors.

4. **Lessons for the others:** Finally, the study's findings can potentially be expanded to broader regions. The power grid of Northern Virginia is currently at its maximum capacity, limiting the potential to physically add new data centres. Other US states, such as Texas, are

speculated to accommodate additional infrastructure buildout e.g., Stargate Project (2025). Incorporating study's findings into accurate demand forecasting (Point 1), allocating financial responsibilities to power grid expansion (Point 2) and ensuring a market design which protects pre-existing businesses and households from price shocks (Point 3), is in direct interest of engaged stakeholders: policymakers, grid operators or data centre owners.

5.2 Limitations

Using the 'synthetic residuals' with a VAR model resembles the intuition behind the proxy-SVAR method (Jordà and Mertens, 2022). Additional theoretical support and evidence of exogeneity would allow expansion of the study's findings into a formal proxy-SVAR experiment. Since this study relies on proxy estimates for data centre activity rather than direct consumption data, any causality claims regarding price and demand impacts are necessarily indirect and, hence, limited. Future research should incorporate further specialised proxies for the activity of data centres such as global internet traffic or usage statistics of popular AI tools e.g., ChatGPT family (OpenAI, 2025). Currently such data is only available through paid services and tends to be disaggregated and unreliable (Shehabi et al., 2024). Furthermore, due to the limitations of this work, future research should incorporate additional theoretical frameworks such as time-varying parameter or Markov-switching models to account for the structural break modelling.

As the study suggests, the data centre industry evolves at a rapid pace. Compute-intensive services such as data storage, cryptocurrencies and LLMs evolve and introduce major efficiency improvements every year (Shehabi et al., 2024). Therefore, existing forecasts and estimates methods may need to be revised and adjusted[13]. New research should account for the evolution of power efficiency of the computing technologies. The Energy Act (2020), which requires data centres to monitor and publish their energy efficiency estimates, offers opportunities for such studies.

Further robustness could be achieved through additional market research on the PJM and data centre industry. Large power-offtakers, such as data centres, tend to limit their exposure to market prices through long-term power contracts (PPAs), which could potentially explain Virginia's commercial electricity trades at, on average, lower prices than industrial or residential sectors (Fig 20). Therefore, data centres may tend to indirectly saturate the regional electricity market, while not participating in the direct price formation. Explaining, verifying and formalising such a mechanism through e.g., a network model[14], presents an interesting avenue for future research.

The study leverages the unique geographic property of Virginia as a quasi-exogenous shock to estimate the effect of data centres on Virginia's commercial electricity consumption and their indirect impact on regional electricity prices. The estimates are counter-verified with official, institutional reviews by EPRI (2024) and Shehabi et al. (2024) with only a 15% mismatch. The analysis provides robust evidence data centres strongly impact the regional electricity prices with an exogenous 1GWh shock leading to, on average, persistent 0.2% price increase. The impact onto long term demand requires additional theory and empirical evidence. Given that data centres are the underlying infrastructure supporting the Internet of Things economy or 'AI' research, their economic, technological and political importance is growing rapidly (e.g., House (2023)). The study seeks to incorporate a novel, economic, perspective on the industry and open new arrays for future research.

[13] For example, Carvallo et al. (2018) suggest *while not meaningless, historical utility demand forecasts consistently overestimate both peak and average demand.*

[14] which incorporates and models an integrated PJM wholesale clearing price mechanism.

References

Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*, 59(2):391–425.

Abadie, A., Diamond, A., and 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.

Abadie, A. and Gardeazabal, J. (2003). The Economic Costs of Conflict: A Case Study of the Basque Country. *American Economic Review*, 93(1):113–132.

Abadie, A. and Vives-i Bastida, J. (2021). Synthetic controls in action. *MIT Economics*.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723.

Andrews, D. W. K. (2003). End-of-sample instability tests. *Econometrica*, 71(6):1661–1694.

Anthropic (2025). Claude (claude 3.7 sonnet) [large language model].

Aurora Energy Research (2024). Aurora report finds northern virginia data center demand could incentivise up to 15 gw of additional natural gas generators by 2030. Accessed: 2025-04-02.

Bai, J. and Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, 66(1):47–78.

BEA (2025). Regional data: Gdp and personal income. Accessed: April 2, 2025.

Bistline, J. E. T., Mehrotra, N. R., and Wolfram, C. (2023). Economic implications of the climate provisions of the inflation reduction act. *Brookings Papers on Economic Activity*, 2023(1):77–182.

Bloomberg L.P. (2025). Bloomberg virginia index, 1/1/2000 to 02/01/2025. Retrieved February 12, 2025, from Bloomberg Terminal.

Borenstein, S., Bushnell, J. B., and Wolak, F. A. (2002). Measuring market inefficiencies in california's restructured wholesale electricity market. *American Economic Review*, 92(5):1376–1405.

Brown, R. L., Durbin, J., and Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society. Series B (Methodological)*, 37(2):149–192.

Carvallo, J. P., Larsen, P. H., Sanstad, A. H., and Goldman, C. A. (2018). Long-term load forecasting accuracy in electric utility integrated resource planning. *Energy Policy*, 118:111–123.

Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, 28(3):591–605.

de Vries, A. (2023). The growing energy footprint of artificial intelligence. *Joule*, 7(10):2191–2194.

DoE (2024). Resources available to support data center electricity needs. Technical report, U.S. Department of Energy.

DoJ (2008). Electricity restructuring: What has worked, what has not, and what is next. Technical Report EAG 08-4, U.S. Department of Justice, Antitrust Division. Main Author: Jeff Lien.

Dominion Energy (2025). Q4 2024 earnings call. Accessed: 2025-04-02.

EIA (2024a). Commercial electricity demand grew fastest in states with rapid computing facility growth. Technical report. U.S. Energy Information Administration.

EIA (2024b). Eia-861m monthly electric power industry report. Short Term Energy Outlook (STEO) - Monthly. U.S. Energy Information Administration, Released: 10/28/2024.

EIA (2024). Virginia was the top net electricity recipient of any state in 2023. U.S. Energy Information Agency.

EIA (2025). EIA-861M Form (STEO EIA-861M) Monthly Electric Power Industry Report. U.S. Energy Information Agency.

EIA (2025a). EIA-861M, HS861M Monthly Electric Power Industry Report. Released: March 25, 2025.

EIA (2025b). U.s. wholesale electricity prices were lower and less volatile in 2024. U.S. Energy Information Administration.

EIA (2025c). Virginia price of natural gas sold to commercial consumers (dollars per thousand cubic feet), series id: N3020va3. U.S. Energy Information Agency 1989–2025 data series.

Energy Act (2020). Energy Act of 2020. Division Z of the Consolidated Appropriations Act, 2021, Public Law No. 116-260.

EPRI (2024). Powering intelligence: Analyzing artificial intelligence and data center energy consumption.

Esmaeili, P. and Rafei, M. (2021). Dynamics analysis of factors affecting electricity consumption fluctuations based on economic conditions: Application of svar and tvp-var models. *Energy*, 226:120340.

Fair, R. C. (2003). Testing for a new economy in the 1990s. Yale ICF Working Paper No. 02-48, Cowles Foundation Discussion Paper No. 1388.

FRED St Louis (2025a). Business applications with planned wages: Total for all naics per state (bawbatotalsava). Federal Reserve Bank of St. Louis.

FRED St Louis (2025b). Exports of non-manufactured commodities (expnonga), monthly, millions of dollars, not seasonally adjusted. Federal Reserve Bank of St. Louis.

FRED St Louis (2025c). Fred-sd: State-level databases for macroeconomic research and forecasting. Federal Reserve Bank of St. Louis.

Granger, C. and Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics*, 2(2):111–120.

Haldrup, N., Nielsen, F. S., and Ørregaard Nielsen, M. (2010). A vector autoregressive model for electricity prices subject to long memory and regime switching. *Energy Economics*, 32(5):1044–1058.

Hill, A. (2021). Price freezes and gas sensitivity: An estimation of the price impact of electricity market restructuring. Accessed April 6, 2025.

Horowitz, M. J. (2007). Changes in electricity demand in the united states from the 1970s to 2003. *The Energy Journal*, 28(3):93–120.

House, T. W. (2023). Fact sheet: One year in, president biden's inflation reduction act is driving historic climate action and investing in america to create good paying jobs and reduce costs. Statement and Releases, August 16th 2023.

FED Philadelphia (2025). State coincident indexes, 1979–2025. Federal Reserve Bank of Philadelphia.

IEA (2024). Electricity 2024: Analysis and forecast to 2026. Technical report. International Energy Agency.

JLARC (2024). Data centers in virginia 2024. Technical Report 598, Joint Legislative Audit and Review Commission. Joint Legislative Audit and Review Commission. Report to the Governor and the General Assembly of Virginia.

Jordà, O. and Mertens, K. (2022). Techniques of empirical econometrics. American Economic Association: Continuing Education, January 8–10, 2022.

Kimley-Horn Staff (2024). Power transmission evaluation of loudoun county, virginia.

Koot, M. and Wijnhoven, F. (2021). Usage impact on data center electricity needs: A system dynamic forecasting model. *Applied Energy*, 291:116798.

Kraft, J. and Kraft, A. (1978). On the relationship between energy and gnp. *The Journal of Energy and Development*, pages 401–403.

Li, R., Woo, C.-K., and Cox, K. (2021). How price-responsive is residential retail electricity demand in the us? *Energy*, 232:120921.

Malmodin, J., Lövehagen, N., Bergmark, P., and Lundén, D. (2024). Ict sector electricity consumption and greenhouse gas emissions – 2020 outcome. *Telecommunications Policy*, 48(3):102701.

Maryland DoC (2020). Data center maryland sales and use tax exemption incentive program. Technical report. Maryland Department of Commerce, Accessed November 17, 2024.

Masanet, E. e. a. (2020). Recalibrating global data center energy-use estimates. *Science*, 367:984–986.

Mytton, D. and Ashtine, M. (2022). Sources of data center energy estimates: A comprehensive review. *Joule Volume 6, Issue 9 Pages 2032-2056 ISSN 2542-4351*.

NOAA (2025). Climate at a glance: Statewide time series. National Centers for Environmental Information, Published March 2025.

OpenAI (2025). Chatgpt (gpt-4o) [large language model].

Ponne, B. (2023). When numbers meet stories – an introduction to the synthetic control method in r. Published on R-bloggers.

Ros, A. J. (2017). An econometric assessment of electricity demand in the united states using utility-specific panel data and the impact of retail competition on prices. *The Energy Journal*, 38(4):73–100.

Shehabi, A., Smith, S. J., Hubbard, A., Newkirk, A., Lei, N., Siddik, M. A. B., Holecek, B., Koomey, J., Masanet, E., and Sartor, D. (2024). United States Data Center Energy Usage Report 2024. Technical Report LBNL-2024, Lawrence Berkeley National Laboratory.

Shehabi, A., Smith, S. J., Masanet, E., and Koomey, J. (2016). United states data center energy usage report. Technical report, Lawrence Berkeley National Laboratory. LBNL1005775.

Shehabi, A., Smith, S. J., Masanet, E., and Koomey, J. (2018). Data center growth in the united states: Decoupling the demand for services from electricity use. *Environmental Research Letters*, 13:124030.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, 48(1):1–48.

Stargate Project (2025). Announcing the stargate project. Accessed: 2025-04-10.

Stock, J. H. and Watson, M. W. (1996). Evidence on structural instability in macroeconomic time series relations. *Journal of Business Economic Statistics*, 14(1):11–30.

TeleGeography (2025). Submarine cable map. Accessed: November 1, 2024.

Trost, R. (1995). An integrated bayesian vector auto regression and error correction model for forecasting electricity consumption and prices. *Journal of Forecasting*.

Turner, M. (2024). Loudoun county, virginia: Data center capital of the world: A strategy for a changing paradigm. Technical report, Loudoun County.

Turner, M. (2025). Loudoun county, virginia: Data center capital of the world - a strategy for a changing paradigm. second edition.

U.S. Census Bureau (2022). American community survey 5-year estimates, median household income by county in the united states and puerto rico. Accessed on November 1, 2024.

Verdecchia, R., Sallou, J., and Cruz, L. (2023). A systematic review of green ai. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 13(4):e1507.

Virginia Economic Review (2019). The dawn of data: How virginia became the data center capital of the world. *Virginia Economic Review*, Fourth Quarter.

6 Appendix

6.1 Electricity markets of Virginia

The data centre industry in Virginia has been consistently growing since early 2000s.

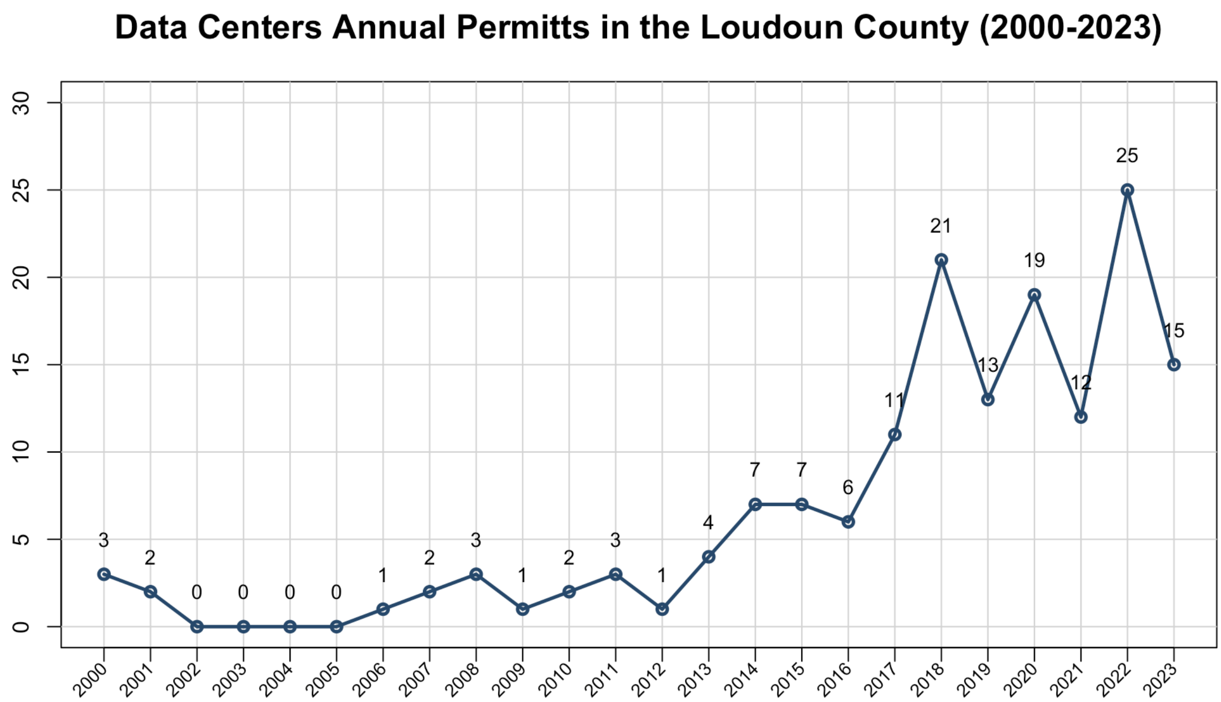


Figure 17: Data centres Permits in the Loudon County. Data Source: (Kimley-Horn Staff, 2024)

Majority of Virginia-based data centres are located within the Loudoun, Prince Williams and Fairfax counties (JLARC, 2024).

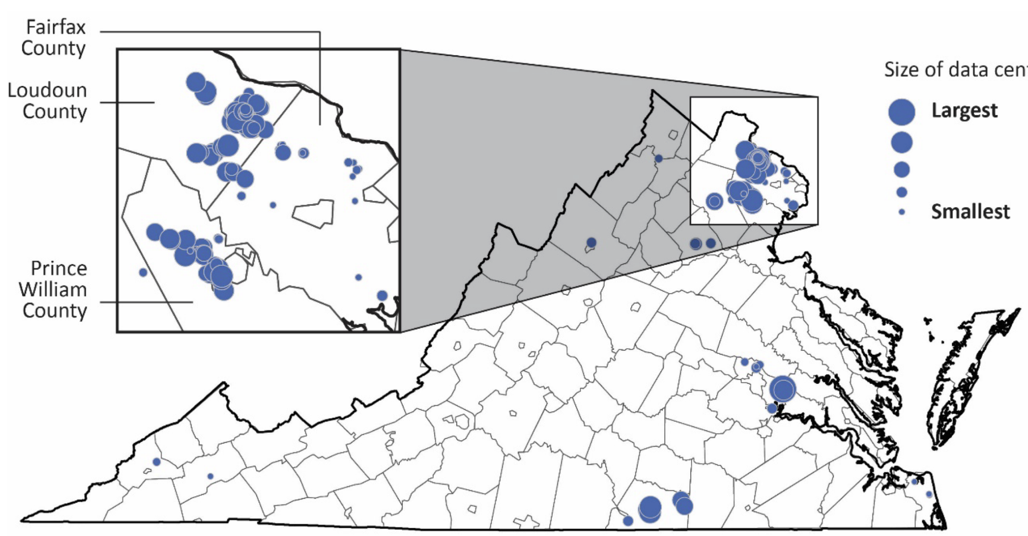


Figure 18: Geographic Distribution of the data centre infrastructure in Virginia. Source: (JLARC, 2024).

The commercial sector has overtaken the residential sector as the leading electricity consumer in state of Virginia.

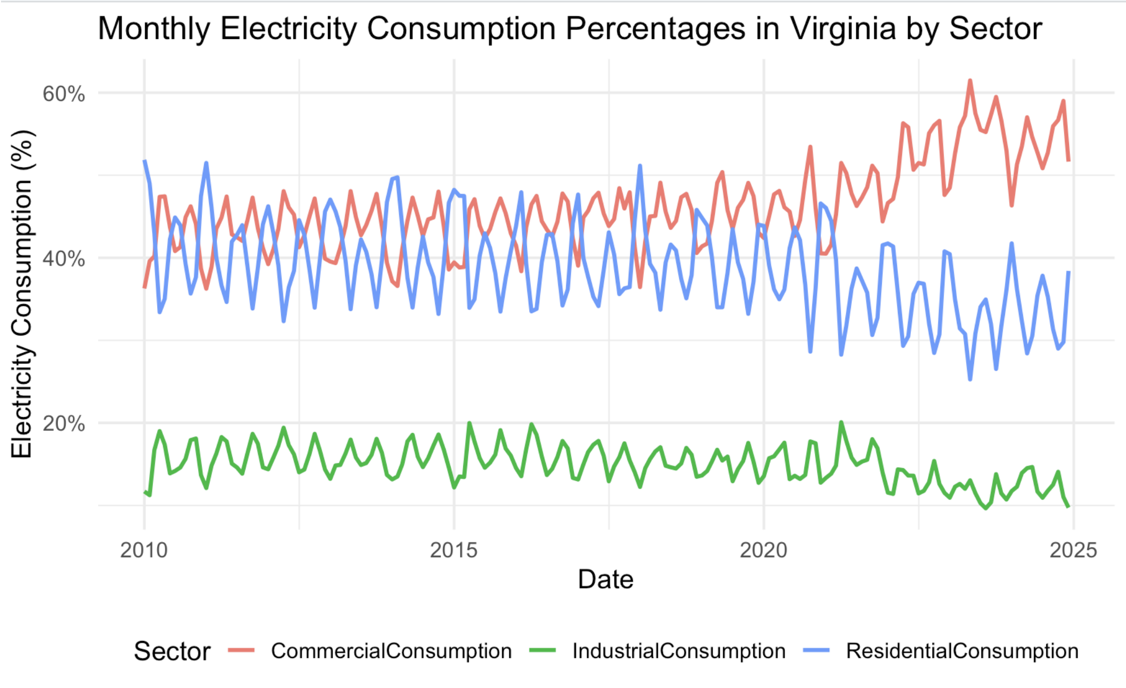


Figure 19: The commercial sector is the leading electricity consumer in Virginia, accounting to over 60% of the entire power usage. Data Source: (EIA, 2024b)

However, this has not transposed into per sector electricity prices. The commercial prices have diverged from its counterparts.

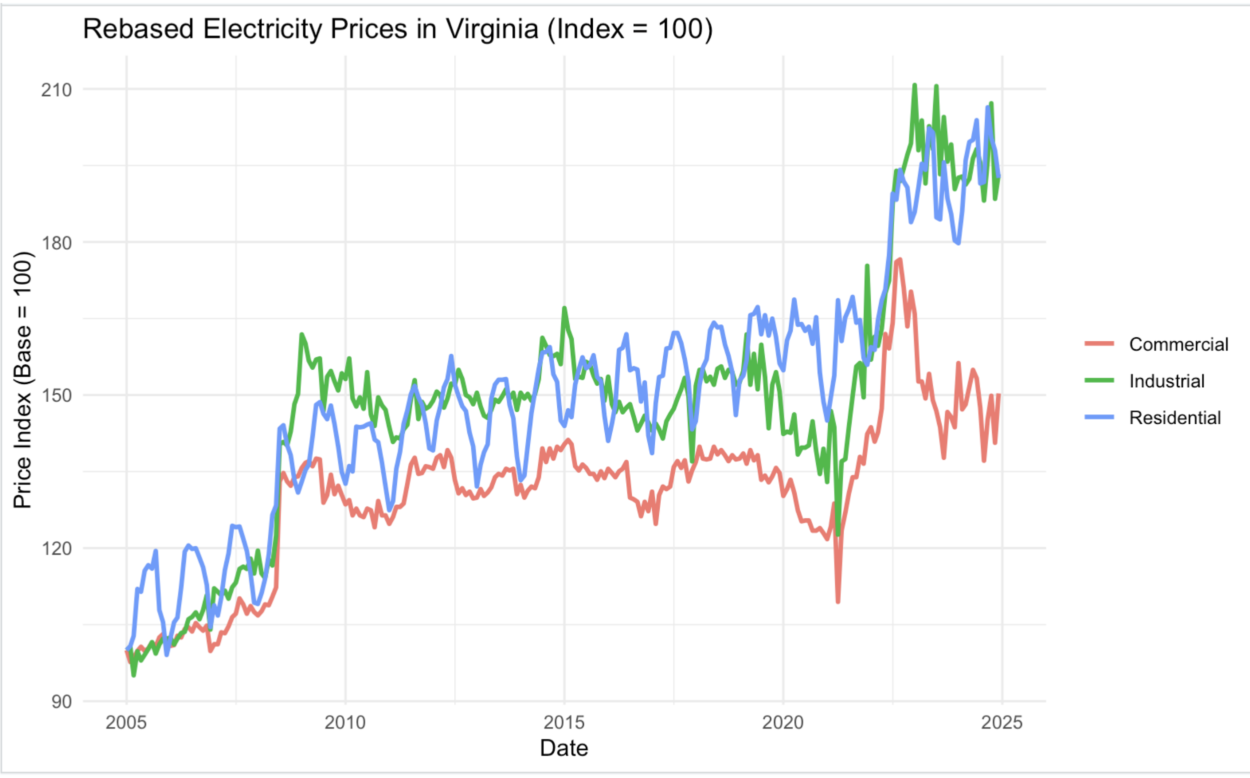


Figure 20: The dynamics of sector-based electricity prices in Virginia has diverged since 2021.

Virginia (8.7 millions inhabitants, 598$ billions GDP) is the biggest net importer of electricity in the entire United States, overcoming California (39 millions inhabitants, 3.9$ trillions GDP) in 2023 (BEA, 2025).

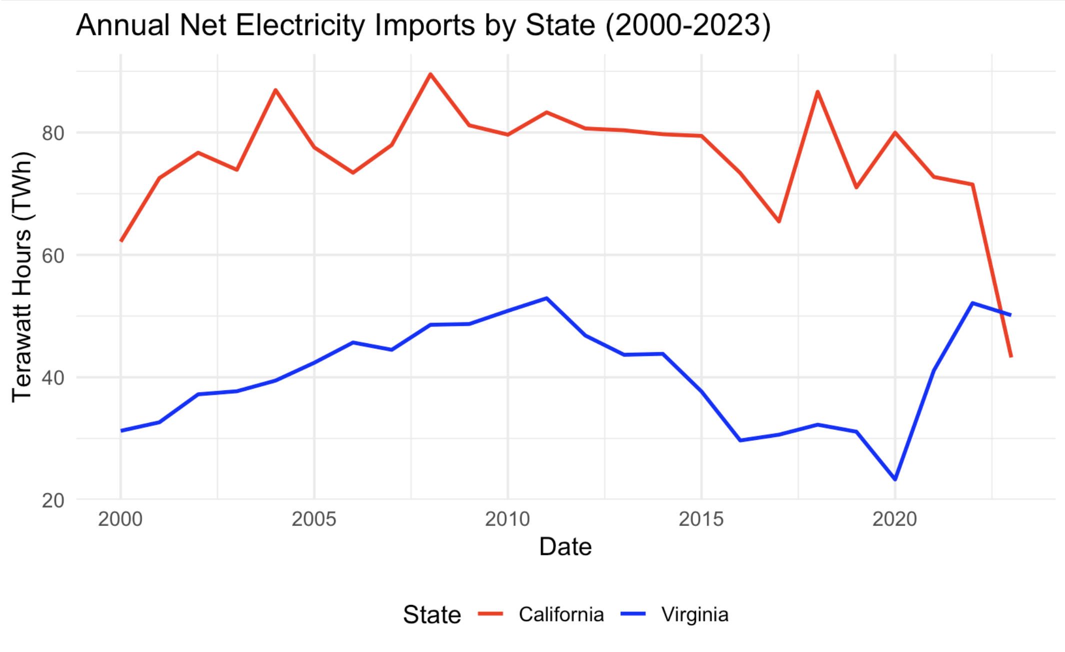


Figure 21: Net interstate electricity imports of California and Virginia (2000-2023) Data Source: (EIA, 2024)

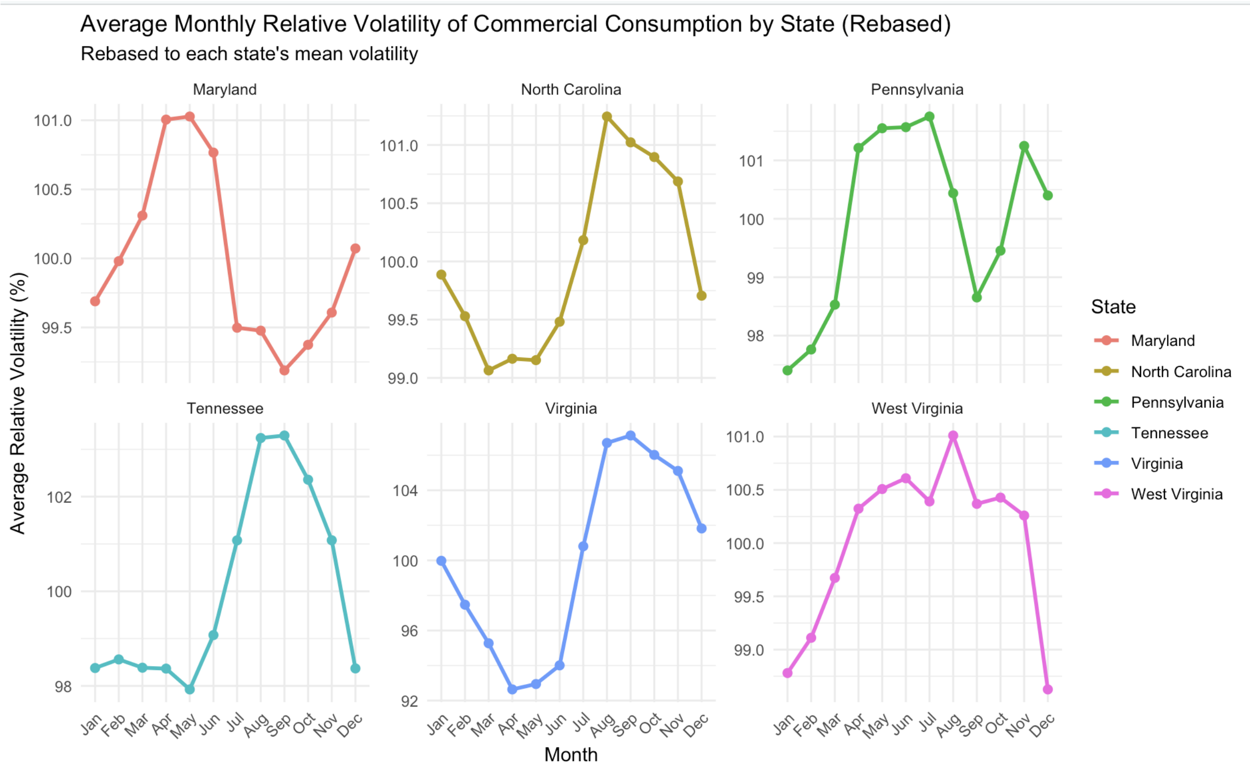


Figure 22: Based on 2008-2024 data, Virginia experiences a greater seasonal volatility of commercial electricity consumption (EIA, 2024b)

Simultaneously, Virginia experiences a greater seasonal volatility of commercial electricity consumption. The energy usage dynamics provided by the Synthetic Control Analysis are in line with evidence provided by (Shehabi et al., 2024).

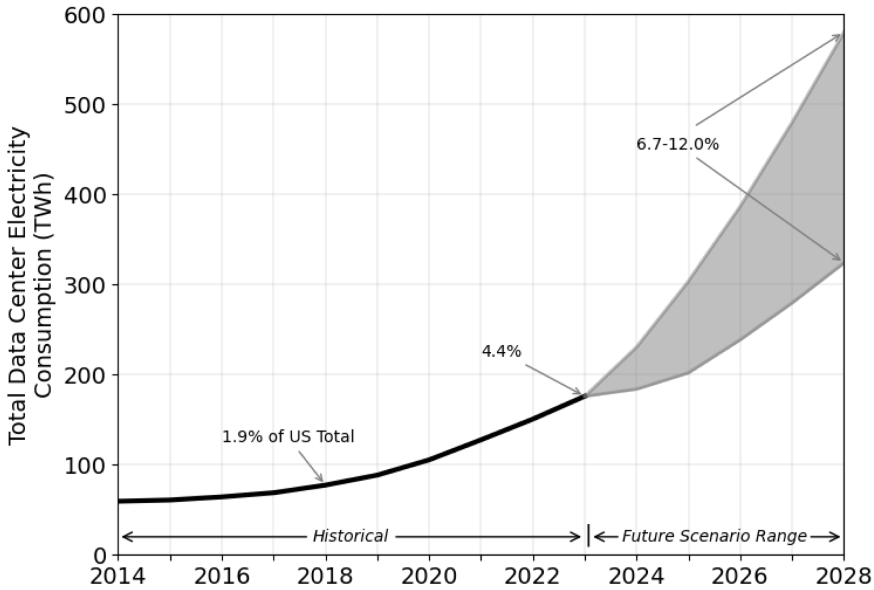


Figure 23: (Shehabi et al., 2024) estimate and forecast a major breakthrough in US electricity consumption from data centres.

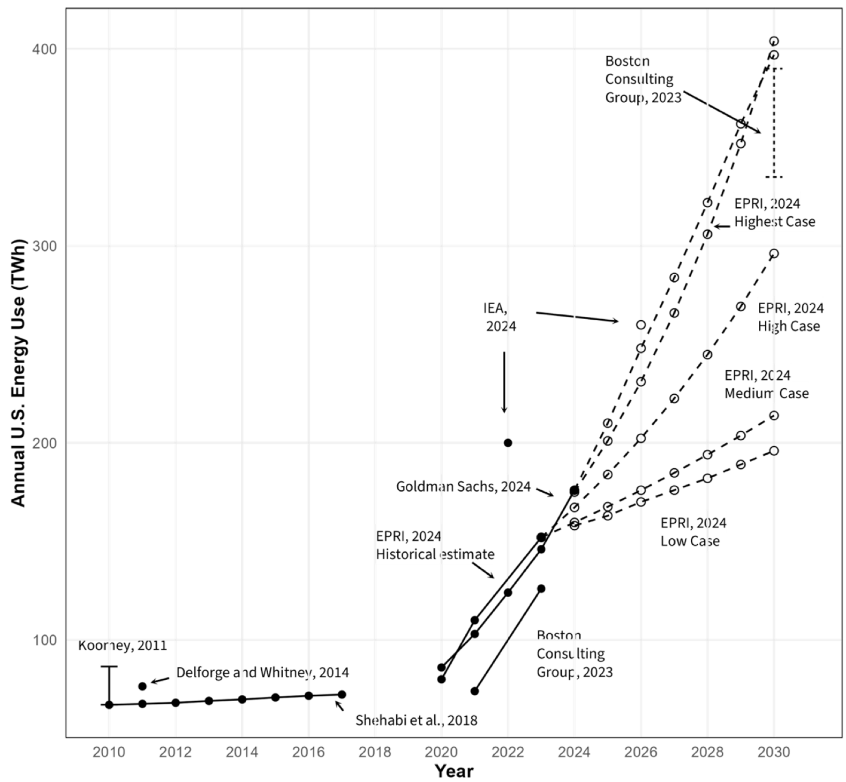


Figure 24: The most recent forecasts of US data centre power consumptions. Source: (Shehabi et al., 2024)

6.2 Methodology

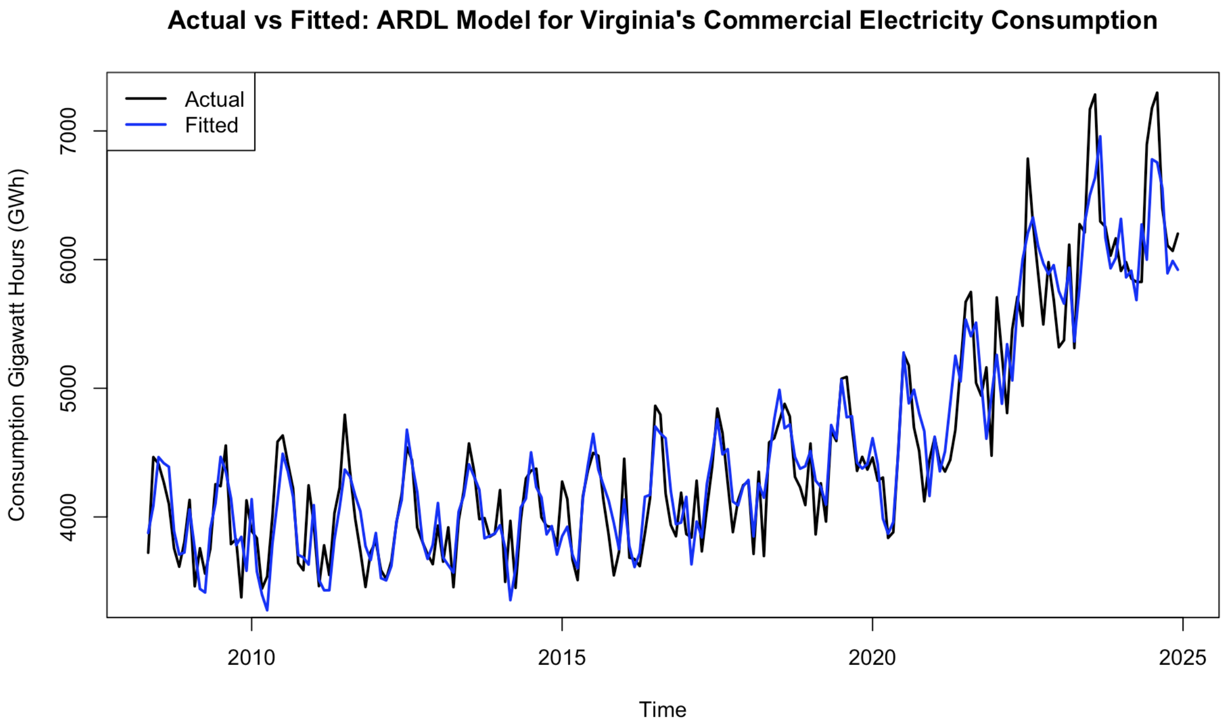
The Chow Test statistic is computed as f-statistic:

$$F = \frac{(SSR_r - (SSR_1 + SSR_2))/k}{(SSR_1 + SSR_2)/(n_1 + n_2 - 2k)}$$

where:

- SSR_r is the sum of squared residuals from the restricted (pooled) model,
- SSR_1 and SSR_2 are the sum of squared residuals from the sub-sample models,
- k is the number of estimated coefficients,
- n_1 and n_2 are the number of observations in each sub-sample.

The ARDL lag selection, which deliberately does not incorporate proxies for data centres, increases residuals over time.



The automatic, optimal weights for Synthetic Control Analysis are:

State	Weight
MD	0.01
NC	0.83
WV	0.01
TN	0.01
PA	0.14

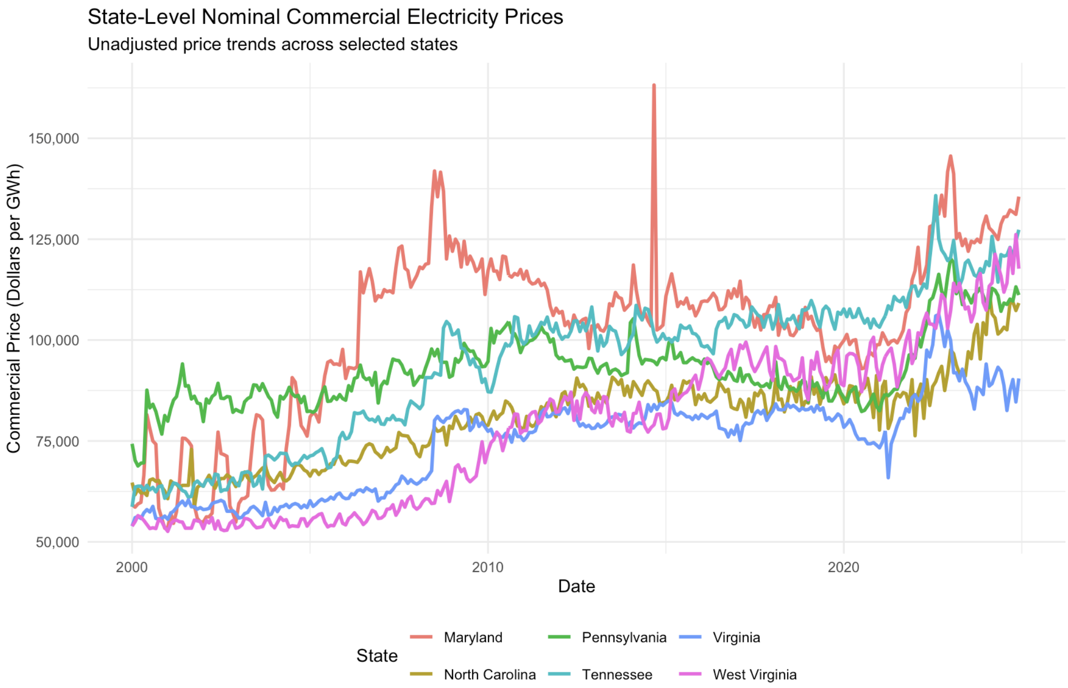
Table 4: Synthetic Control Weights for Virginia

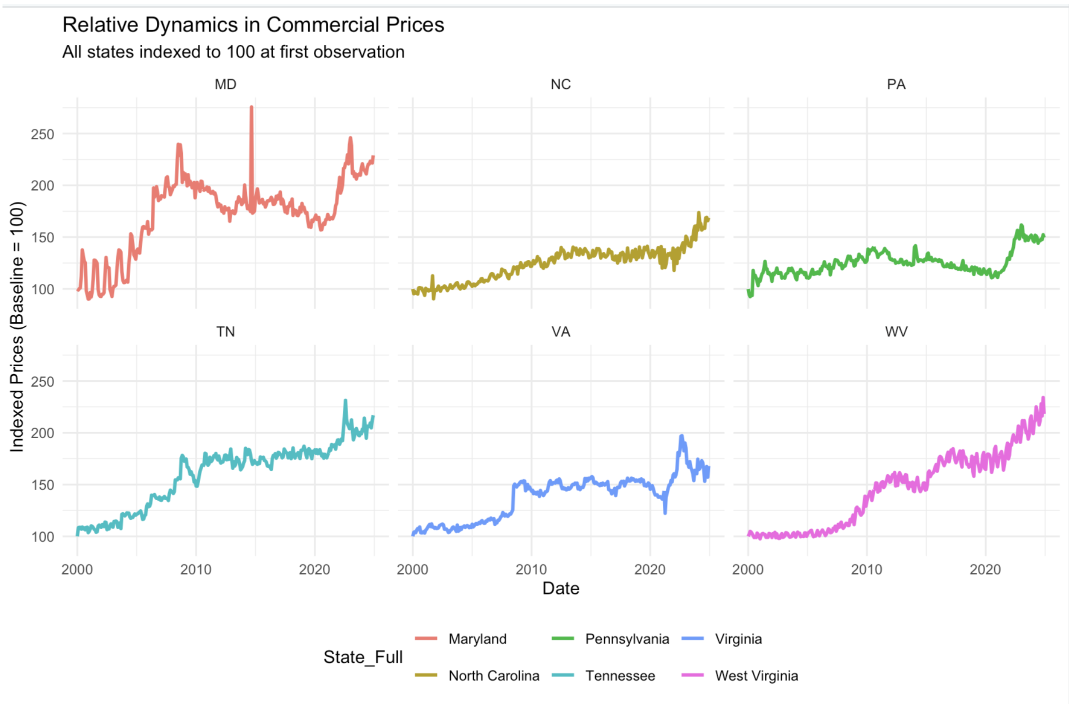
The Johansen Trace Test Results reveal no evidence for co-integration:

	Rank	Test_Statistic	P-10 Critical.Value	P-5 Critical.Value	P-1 Critical.Value
r <= 4 |	0	5.68	7.52	9.24	12.97
r <= 3 |	1	17.86	17.85	19.96	24.60
r <= 2 |	2	35.80	32.00	34.91	41.07
r <= 1 |	3	74.30	49.65	53.12	60.16
r = 0 |	4	172.61	71.86	76.07	84.45

Table 5: Johansen Trace Test Results: No Cointegration Detected

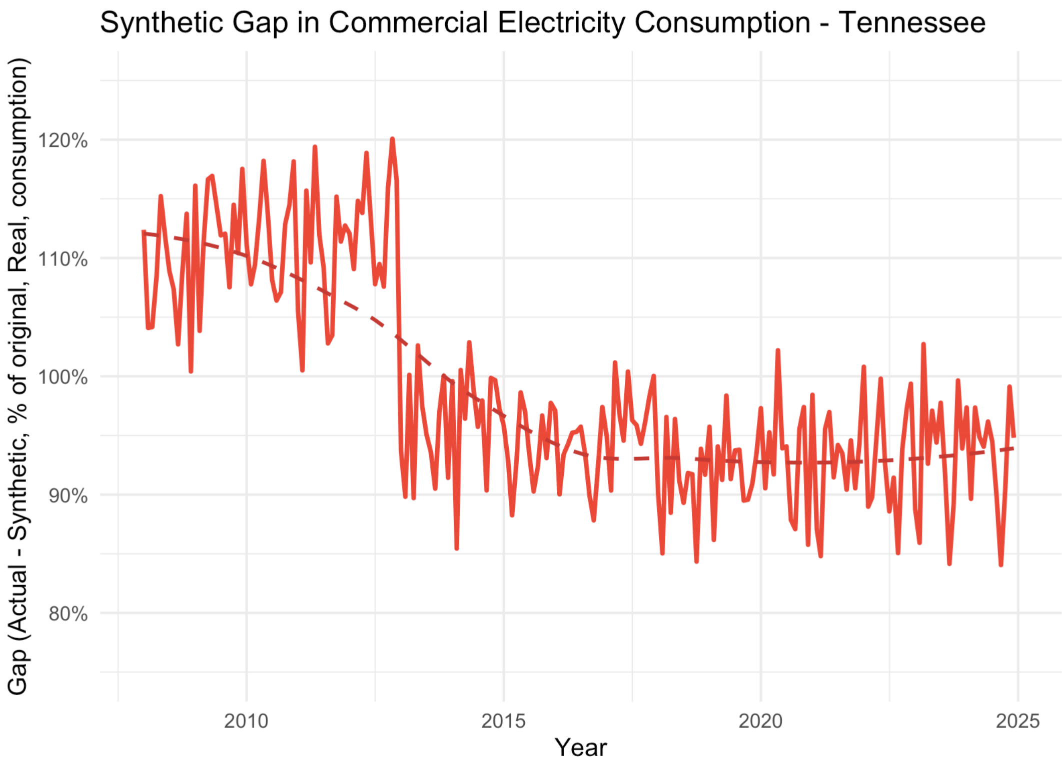
Remarkably, Virginia's commercial electricity prices are both nominally and relatively lower than of neighbouring states.

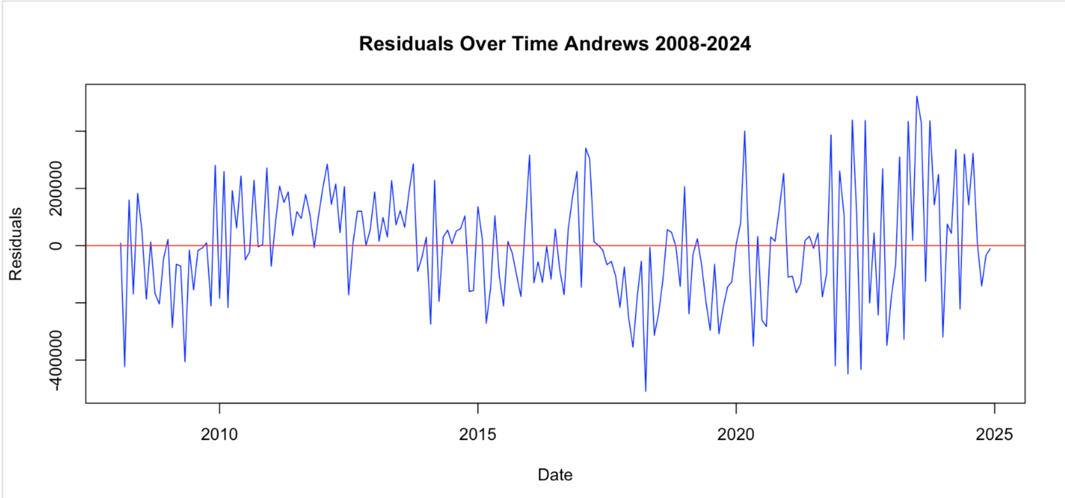


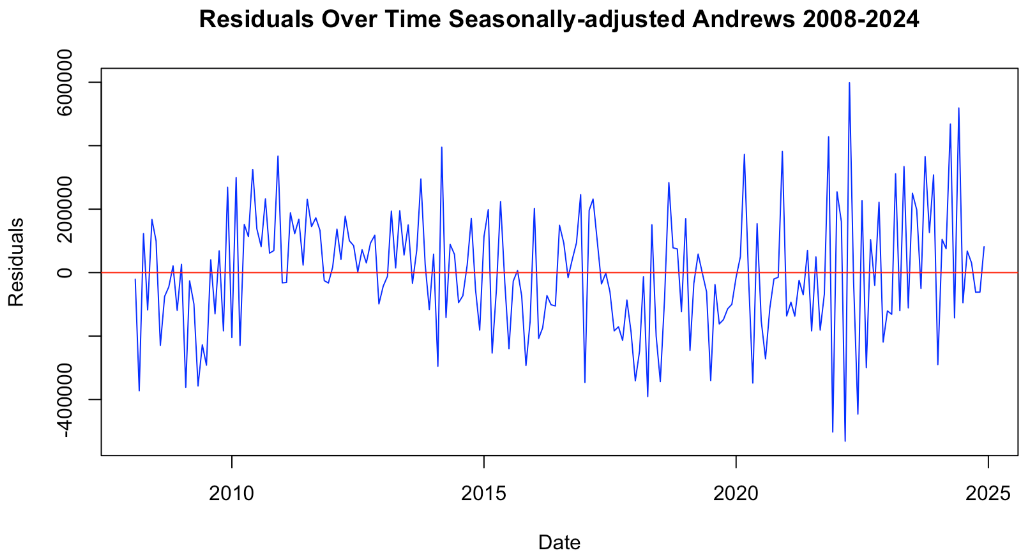


6.3 Results

I perform a 'Placebo' Test for Synthetic Control Analysis for the state of Tennessee. The results demonstrate no substantial divergence between *Synthetic* and *Real* electricity consumption - the Synthetic Gap varies within a 10% bound between *Real* and *Synthetic* gaps, which increases robustness of the original SCA for Virginia.







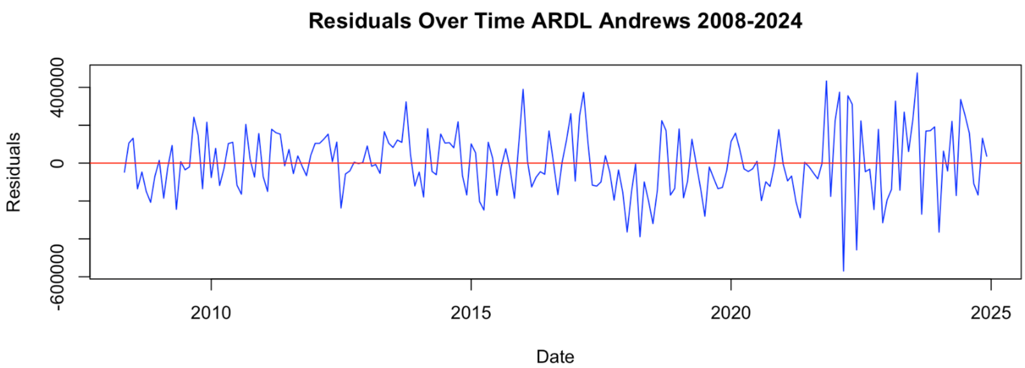


Figure 25: The residuals of several linear models incorporating the Andrews Test tend to increase after 2020 which indicates some omitted variable bias.