Sequential Investment Under Uncertainty

Stein-Erik Fleten[1] and Carl J. Ullrich[2†]

December 26, 2025

Abstract

We investigate sequential investment in electric power generators in PJM. We classify proposed generators into four stages, Planning, Construction, Indefinitely Postponed, and Canceled. We find that duration in Planning, project size, and capacity (RPM) prices have stable and meaningful effects on transitions: longer time in Planning reduces the probability of Construction and increases the probability of Cancellation; larger projects move slowly; and higher capacity prices shift probability mass from Planning to Construction without increasing Cancellation. We contribute a novel method to quantify resistance (***Headwinds***) and uncertainty (***Turbulence***) in the planning process. We show that an increase in uncertainty (***Turbulence***) is associated with a lower probability that a proposed fossil fuel generator remains in Planning and a higher probability of entering Construction.

JEL Classification Codes: D81, E22, G31, L51, L94, AQ41, Q48.

Key Words: Sequential investment, regulatory uncertainty, real options, electricity market restructuring.

[†]Corresponding author. E-mail: ullriccj@jmu.edu.

[1]Norwegian University of Science and Technology, Department of Industrial Economics and Technology Management, NO-7491, Trondheim, Norway.

[2]James Madison University, Department of Finance and Business Law, MSC0209, Harrisonburg, VA, 22807

1 Introduction

We study sequential investment in electric power generators in the PJM Interconnection L.L.C. (PJM). Our primary data source is Energy Information Administration (EIA) form 860 which includes data for all electric power generators–existing, planned, and canceled–in the U.S. Importantly for our study, generators that are planned then canceled remain in the database.

Planning and constructing new electric power generators is a multistage, capital-intensive, mostly irreversible process. Real options theory suggests that uncertainty can affect investment timing at any stage of the process (Roberts and Weitzman, 1981). Decision makers may wait for better information (for uncertainty resolution) before exercising the option to proceed to the next stage of the process. Walls et al. (2007) point out that generators are abandoned sequentially until only projects that will be completed remain.

The theory of irreversible investment (McDonald and Siegel, 1986; Dixit and Pindyck, 1994) predicts that uncertainty and discretion over investment timing lead to deferral of investment compared to now-or-never alternatives. Empirical work tends to support this theory (Quigg, 1993; Paddock et al., 1988; Bulan, 2005; Bulan et al., 2009; Moel and Tufano, 2002; Fleten et al., 2016, 2017), although there is also evidence that when government licenses to invest have been granted, the projects tend to go on to completion (Somerville, 2001; Linnerud et al., 2014).

Also in theoretical work where investments occur in stages, increased uncertainty tends to delay investment (Majd and Pindyck, 1987; Gollier et al., 2005; Siddiqui and Fleten, 2010; Chronopoulos et al., 2016), as more information can be gained by waiting. However, uncertainty can also accelerate staged or exploratory investment, as firms value the information gained or the flexibility retained (Bar-Ilan and Strange, 1998; Huchzermeier and Loch, 2001; Chronopoulos et al., 2017). Empirical work finds both hastening (Favero et al., 1994; Marmer and Slade, 2018) and delay (Moel and Tufano, 2002; Bulan et al., 2009; Kellogg, 2014; Fleten et al., 2017). Our empirical setting, with multi-staged electricity generation investment, allows us to test these predictions.

The electricity industry is characterized by long lead times, capital intensity and long lifetimes of the assets, making the theory of irreversible investment especially relevant (Fleten et al., 2007; Ishii and Yan, 2011; Walls et al., 2007; Fabrizio, 2013; Fleten et al., 2024). In this industry, some assets (combustion turbines) have operational flexibility, while others (solar and wind) are running more or less when possible. Näsäkkälä and Fleten (2005) show theoretically that increased volatility affects these two groups differently.

A related issue is how firms form expectations regarding future profitability and profitability uncertainty (Fleten et al., 2017). Structural approaches seem useful for eliciting such deep aspects of decision making (Cook and Lin Lawell, 2020; Elliott, 2022; Fleten et al., 2020; Çam et al., 2022). The empirical fit of these models supports the idea that electricity investors behave as if they anticipate future states rather than reacting only to current profitability.

We use the EIA 860 variable **STATUS** to classify each proposed generator into a **STAGE** as in Table 1 below. We have 411 proposed generator groups[1] and 870 total generator group-year observations in our 2008–2023 sample.

Because the EIA860 STATUS code definitions change in 2016 (see Table 1), we also consider a subsample from 2016–2023. in the subsample we have 293 generator groups and 562 generator group-year observations.

We introduce novel measures of resistance, ***Headwinds***, and uncertainty, ***Turbulence***, in the planning environment. *Headwinds* is the technology-year average time delay which projects experience in planned operational dates, while *Turbulence* is the standard deviation of delays. We categorize generators into one of two technology types, either fossil fuel or renewable. Fossil fuel includes generators which use coal, natural gas, and/or oil. Renewable generators use sun, wind, and/or water. Intuitively, stronger ***Headwinds*** indicate more impediments in the planning environment and can reflect factors such as regulatory delays, changes in profitability reassessments, environmental compliance issues, and technological risks. Unlike traditional uncertainty measures such as Baker et al. (2016) and its variants, ***Turbulence*** captures uncertainty comprehensively, is specific to the investments at hand, and is based upon revealed preferences.[2]

We recognize that endogeneity is a potential issue as the factors which cause a revision in in-service dates might also affect the transition decision. We take steps to mitigate these concerns and discuss the issue in detail below. Specifically, we calculate our new indices using nationwide data, then recalculate the indices using only data not in our study sample.[3] Still, we are careful to interpret our results as associational, not strictly causal.

We employ a multinomial logit model for sequential investment decisions.[4] Our empirical results reveal several robust patterns.

Duration, i.e., the length of time a generator has been in the PLG stage, is a strong predictor of transitions. Projects that have been in the Planning (PLG) stage for only a short period are likely to remain in PLG. After roughly three to four years, additional time in PLG increases the probability of exiting planning. In the full sample this exit occurs through the combined IDP/CNL category, while in the 2016–2023 subsample it occurs primarily through cancellation (CNL).

Project scale matters. Larger projects (higher Nameplate capacity) and larger multi-unit project groups (larger GroupSize) are consistently less likely to advance into construction, reflecting greater irreversibility or complexity. Multi-unit groups are also substantially less likely to be canceled in the subsample.

[1]We refer to sets of identical proposed generators as generator groups. Identical generators are located at the same plant, have the same prime mover, the same capacity, the same primary and secondary fuels, and the same sequence of progression through the **STAGE**s defined in Table 1. In the following we use the term *generator* to refer to the entire group, unless otherwise stated. We create a new variable *GroupSize* which is equal to the number of generators in a group. Singleton generators have *GroupSize* = 1.

[2]Both *Headwinds* and *Turbulence* carry unique information, but due to the sparseness of our data, they have significant overlap. In the regression analysis we use a residualized version of *Turbulence* that is orthogonal to *Headwinds*; see Section 2.3.5 for details.

[3]The sample used to create the indices and the sample used in the regressions are completely separate.

[4]The source code used in the analysis is written in **STATA** and **Python**, and is available from the corresponding author.

Capacity prices (CapPmt) are important. Higher capacity prices robustly increase the probability of moving out of Planning (PLG) and into Construction, while leaving Cancellation (CNL) probabilities unchanged—consistent with RPM providing an investment incentive.

Uncertainty has asymmetric effects across technologies. In the full sample (2008–2023), the effects of *Headwinds* and *Turbulence* are economically small. In the 2016–2023 subsample sparse IDP/CNL transitions make them difficult to estimate precisely. Higher Turbulence accelerates fossil transitions into Construction, whereas renewable responses are weaker and often statistically insignificant. We hypothesize that this behavior reflects the fact combustion turbines (the majority of our fossil fuel sample) are call options. Uncertainty can therefore increase the generator's value and can accelerate the sequential investment process. Renewable projects, lacking similar operational flexibility, do not exhibit such behavior.

The remainder of this paper is structured as follows. In Section 2 we detail the data. Section 3 contains the regression results and a discussion thereof. Section 4 examines marginal effects. Section 5 concludes.

2 Data

Primary data sources are EIA 860 for data on proposed generators and PJM for capacity prices.

2.1 Generators and Status Changes

Each year firms in the United States must report on EIA 860 the status of proposed (and existing) electric power generators. EIA 860 is the source of the data which makes the analysis possible. Table 4 (in Appendix A) lists the variables of interest we take from EIA 860. We undertook extensive data cleaning (see Appendix A) to ensure that we could track generators consistently through time, consolidate generators appropriately, and, correct mistakes in the recorded data. This reduces measurement error in our key variables and assures that our results are not artifacts of data inconsistencies.

The EIA 860 variable **STATUS** is key, as it reveals investment decisions made each year. We obtain from EIA 860 the yearly **STATUS** codes of proposed generators in the PJM footprint. We use the **STATUS** code to define the **STAGE** as in Table 1.[5]

Prior to 2016, **STATUS** code CN *(canceled, previously planned)* did not exist. The 2015 EIA 860 documentation describes **STATUS** code IP as *(Planned new generator canceled, Indefinitely Postponed, or no longer in resource plan).*

The 2016 EIA 860 documentation describes both **STATUS** code IP *(Planned new Indefinitely Postponed, or no longer in resource plan)* and **STATUS** code CN *(canceled previously reported as "planned").*

[5]We drop all regulated generators. Regulated generators almost never are canceled. We have a total of 65 proposed generator group-year observations for regulated generators. (All of which are omitted from the full and subsamples in the analysis below.) Of those, only one transitions to either IDP or CNL.

We consider two sample periods. For the full sample analysis we recode **STAGE** CNL (from the 2016–2023 data) to be **STAGE** IDP, thereby ensuring a consistency throughout the sample.

In the 2016–2023 subsample (293 generator groups and 562 generator group-year observations) we consider **STAGE**s IDP and CNL separately.

Table 1: EIA 860 ***STATUS*** *codes for proposed generators and the corresponding* ***STAGE****s.*

STAGE	EIA 860 **STATUS**	**Description**
PLG (Planning)	P	Planned, no regulatory approval
	L	Planned, regulatory approvals pending
	T	Planned, regulatory approvals received
CON (Construction)	U	Planned, under Construction, less than 50%
	V	Planned, under Construction, more than 50%
	TS	Planned, Construction complete but not in operation
OPR (Operational)	OP	Existing, operating
IDP (Indefinitely Postponed)	IP	Indefinitely postponed/canceled (2008-2015)
	IP	Indefinitely postponed (2016–2023)
CNL (Canceled)	CN	Canceled (2016–2023)

Consider a generator in the Planning **STAGE** (PLG) in the current period t. In the next period $t+1$, that generator can be in any of the **STAGE**s PLG, CON (Under Construction), IDP (Indefinitely Postponed), or CNL (Canceled, in the 2016–2023 subsample).[6]

Figure 1 presents the observations of transitions for the full sample (2008–2023). Figure 2 includes only the subsample (2016–2023).

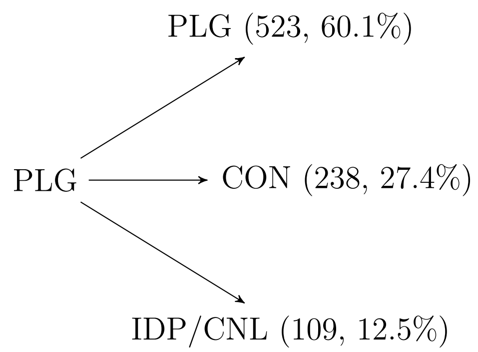

Figure 1: Transitions from Stage PLG in period t to Stages PLG, CON, and IDP/CNL in period $t+1$. Numbers in parentheses show counts and percentages of total transitions over the full sample period (2008–2023).

[6]Because the data are reported at the annual frequency, it can happen that a generator which is in PLG in year t moves all the way through CON within year t and shows up in OPR in year $t+1$. We recode these observations to be in CON in year $t+1$.

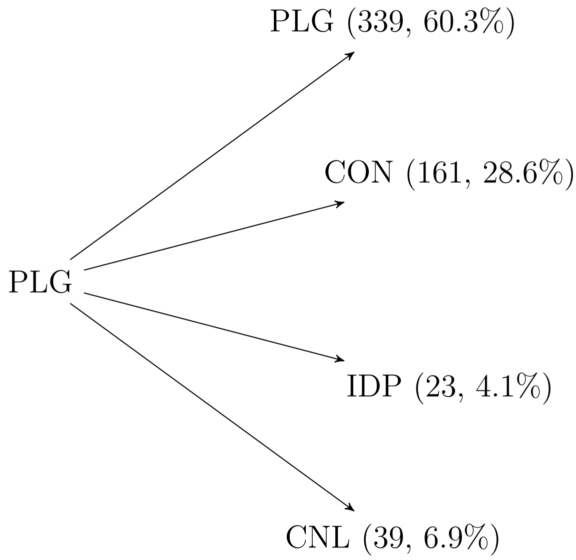

Figure 2: Transitions from Stage PLG in period t to Stages PLG, CON, IDP, and CNL in period $t+1$. Numbers in parentheses show counts and percentages of total transitions over the subsample (2016–2023).

2.2 Capacity Prices

We take zonal capacity prices from the PJM RPM three-year ahead base-residual auction.

2.3 *Headwinds* and *Turbulence*

Beginning with raw data from EIA 860, we construct measures of resistance, which we label ***Headwinds***, and uncertainty, which we label ***Turbulence***, in the planning process. ***Headwinds*** quantifies how strongly the planning environment pushes projects backward (delays) and ***Turbulence*** quantifies variability in those delays.

Specifically we calculate for each generator i in year t the revision (*DiffYear*) in its expected operational date.

$$DiffYear_{i,t} = \left(cuyr_{i,t} + \frac{cumn_{i,t} - 1}{12} \right) - \left(efyr_{i,t} + \frac{efmn_{i,t} - 1}{12} \right), \tag{1}$$

where

- $efmn_{i,t}$: Effective Month - original in-service month.
- $efyr_{i,t}$: Effective Year - original in-service year.
- $cumn_{i,t}$: Current Month - most recently updated in-service month.
- $cuyr_{i,t}$: Current Year - most recently updated in-service year.

DiffYear equals the change in the expected in-service date, measured in decimal years. A positive value indicates delay relative to the original plan, while a negative value represents an acceleration.

2.3.1 Technology Groups

We classify each generator into one of two technology groups:

$$k \in \{1, 2\} \quad \text{(fossil fuel, renewable)}.$$

We define fossil fuel generators to be those generators which use coal, natural gas, or oil. We define renewable generators to be those generators which use sun, wind, or water. We calculate indices separately for each fossil fuel and renewable generators in each year t.

2.3.2 Weighting at the Project Level

Each observation corresponds to a group of identical generators[7] and each group is weighted by its aggregate nameplate capacity. We normalize Weights such that their mean equals 1 within each technology–year cell.

Let w_i denote the normalized weight for observation i.

2.3.3 ***Headwinds***: Mean Delay within a Technology Group

For each technology k and year t, we compute the weighted mean delay:

$$\mathit{Headwinds}_{k,t} = \frac{\sum_{i \in I_{k,t}} w_i \, \mathit{DiffYear}_{i,t}}{\sum_{i \in I_{k,t}} w_i}, \tag{2}$$

where $I_{k,t}$ is the set of all generator groups of technology k observed in year t.

Headwinds summarizes the average delay faced by generators of technology k in year t. Each block inherits its technology's *Headwinds* value for its year. *Headwinds* captures the net effect of the planning environment on a specific technology—including for example regulatory delays, interconnection queues, changes in expected profitability, and any other shock which affects the planning process.

2.3.4 ***Turbulence***: Within-Group Dispersion of Delays

While *Headwinds* measures the length of delays, *Turbulence* measures the cross-sectional variation in delays. For each technology k and year t, we compute the weighted standard deviation of delay:

$$\mathit{Turbulence}_{k,t} = \sqrt{\frac{\sum_{i \in I_{k,t}} w_i \left(\mathit{DiffYear}_{i,t} - \mathit{Headwinds}_{k,t}\right)^2}{\sum_{i \in I_{k,t}} w_i}}. \tag{3}$$

Turbulence captures the degree of heterogeneity in project-level revisions for that technology–year. High values of *Turbulence* indicates variability across projects–some are moving quickly[8] while others are significantly delayed, consistent with a more uncertain or unstable planning environment.

[7] The unit of economic decision making, as described above in FN 1.

[8] It can happen that the projected in-service date decreases, i.e., hastening commencement. In this case ***Headwinds*** become ***Tailwinds***.

Turbulence measures uncertainty directly from revealed project timelines. Unlike binary policy dummies or qualitative narratives, *Turbulence* reflects firms' actual revisions of their own expectations.

2.3.5 Summary

In the two-technology framework, we compute both *Headwinds* and *Turbulence* separately for each technology and each year, using weighted project-level data. This approach preserves the full technological heterogeneity of the development process and provides inputs that are directly aligned with the structure of the sequential investment regressions that follow.

Residualization of *Turbulence* for Regression Analysis. *Headwinds* and *Turbulence* are mechanically correlated: years with large average delays tend also to have large cross-sectional dispersion of delays. To reduce this mechanical correlation and avoid collinearity in the regressions, we orthogonalize *Turbulence* with respect to *Headwinds* before constructing the regressors in Section 3. Specifically, at the technology–year level we run a simple pooled regression of the form

$$Turbulence_{k,t} = \alpha + \beta\, Headwinds_{k,t} + \varepsilon_{k,t}, \tag{4}$$

and define

$$Turbulence^{\mathrm{resid}}_{k,t} \equiv \varepsilon_{k,t}. \tag{5}$$

The *Turbulence* regressors that enter Specification (4) (equation 10) are built from $Turbulence^{\mathrm{resid}}_{k,t}$ rather than the raw $Turbulence_{k,t}$. This ensures that, within each technology group, the *Turbulence* terms capture variation in cross-sectional dispersion that is not already explained by the average delay level (*Headwinds*).

2.3.6 Endogeneity

Endogeneity is a potential concern for the generator-specific variable *DiffYear* and by extension for *Headwinds* and *Turbulence*.

Endogeneity in this context could arise through three main channels: (i) **measurement endogeneity**, if the in-service date is revised as a direct result of the transition decision; (ii) **simultaneity**, if both the transition decision and the in-service date revision are influenced by the same contemporaneous factors; and (iii) **omitted variable bias**, if unobserved characteristics jointly affect both variables. We control for (i) and (ii).[9]

Measurement endogeneity can arise if the transition decision directly influences the updated in-service date. We explicitly control for this direct mechanical connection by calculating the *Headwinds* and *Turbulence* from those states which are not in our regression sample.[10] That is, the sample used to calculate *Headwinds* and *Turbulence* is

[9] As of the time of this draft, we are working on efforts to mitigate (iii).

[10] The 14 states omitted from the calculations of *Headwinds* and *Turbulence* are DC, DE, IL, IN, KY, MD, MI, NC, NJ, OH, PA, TN, VA, WV.

completely separate from the sample we use in the regressions below, in which *Headwinds* and *Turbulence* are explanatory variables.

We mitigate the simultaneity issue by using lagged uncertainty measures. We use year t values of *Headwinds* and *Turbulence* to explain transition decision made in year $t+1$.

2.4 Bayesian Partial Pooling of *Headwinds* and *Turbulence*

The raw technology–level indices described above (*Headwinds* and *Turbulence*) are weighted means and weighted standard deviations of in-service date revisions (*DiffYear*), computed separately for each technology k and year t. Many technology–year cells contain relatively few proposed projects, making the sample means and sample variances noisy. This produces two empirical problems:

1. *Instability*: raw indices fluctuate sharply in years with few observations.[11]
2. *Mechanical correlation*: years with large positive delay levels also tend to display large cross-sectional dispersion, even when the true volatility is modest.

To address these issues, we replace raw sample statistics with estimates obtained from a Bayesian hierarchical model fitted to standardized group-level data. We implement the model in Stan and perform partial pooling across years within each technology, producing stabilized estimates of both delay levels and delay dispersion while preserving genuine time variation and technology heterogeneity. A description of the pooling process and a visual comparison of raw versus pooled indices for each technology appears in Appendix B (Figures 9 through 12).

Finally, we use the **Pooled–PJM states omitted** directly in the regressions below. All PJM generators enter the regression but never enter the index construction; the indices use only non-PJM data. Their smoothed structure produces economically interpretable patterns that track well-documented events in electricity markets (e.g. the 2005–2009 coal cancellations, the wind and solar PTC cycles, the 2018 solar tariff, the 2020 COVID disruptions, and the 2022 fuel-price shock). See Table 5. This link between block-level revisions and macro-level policy and market events is precisely what motivates the index construction and why these indices constitute a contribution of the paper.

3 Regression Results: Coefficient Estimates

In this section we employ multinomial logistic regressions to examine the drivers of sequential investment.

3.1 Model Specifications

The regression specification for the full sample follows.[12]

[11] This effect shows up clearly in the early years (2001–2005) for renewables, in Figures 11 and 12.

[12] The outcome variable $\mathbf{STAGE}_{i,t+1}$ denotes the stage of generator i in year $t+1$. In the full sample analysis, $m \in \{\text{PLG, CON, IDP/CNL}\}$. In the 2016–2023 subsample, $m \in \{\text{PLG, CON, IDP, CNL}\}$.

$$\Pr\Big(\mathbf{STAGE}_{i,t+1} = m \mid \mathbf{Transition}_{i,t}\Big) = \frac{\exp\Big(\mathbf{Transition}_{i,t}\beta_m\Big)}{\sum_{j=1}^{M}\exp\Big(\mathbf{Transition}_{i,t}\beta_j\Big)} \tag{6}$$

Specification 1:

$$\begin{aligned}\mathbf{Transition}_{i,t} = {} & \beta_0 + \beta_1\,\mathit{Nameplate}_i + \beta_2\,\mathit{Duration}_{i,t} + \beta_3\,\mathit{CapPmt}_{t+1,z} + \beta_4\,\mathit{GroupSize}_{i,t} \\ & + \beta_5\,\mathit{Renewable}_i + \beta_6\,\mathit{ZonePlannedCap}_{t,z}\end{aligned} \tag{7}$$

Specification 2:

$$\begin{aligned}\mathbf{Transition}_{i,t} = {} & \beta_0 + \beta_1\,\mathit{Nameplate}_i + \beta_2\,\mathit{Duration}_{i,t} + \beta_3\,\mathit{CapPmt}_{t+1,z} + \beta_4\,\mathit{GroupSize}_{i,t} \\ & + \beta_5\,\mathit{Renewable}_i + \beta_6\,\mathit{ZonePlannedCap}_{t,z} \\ & + \beta_7\,\mathit{Headwinds}_{ff,t} + \beta_8\,\mathit{Headwinds}_{ren,t}\end{aligned} \tag{8}$$

Specification 3

$$\begin{aligned}\mathbf{Transition}_{i,t} = {} & \beta_0 + \beta_1\,\mathit{Nameplate}_i + \beta_2\,\mathit{Duration}_{i,t} + \beta_3\,\mathit{CapPmt}_{t+1,z} + \beta_4\,\mathit{GroupSize}_{i,t} \\ & + \beta_5\,\mathit{Renewable}_i + \beta_6\,\mathit{ZonePlannedCap}_{t,z} \\ & + \beta_7\,\mathit{Turbulence}_{ff,t} + \beta_8\,\mathit{Turbulence}_{ren,t}\end{aligned} \tag{9}$$

Specification 4

$$\begin{aligned}\mathbf{Transition}_{i,t} = {} & \beta_0 + \beta_1\,\mathit{Nameplate}_i + \beta_2\,\mathit{Duration}_{i,t} + \beta_3\,\mathit{CapPmt}_{t+1,z} + \beta_4\,\mathit{GroupSize}_{i,t} \\ & + \beta_5\,\mathit{Renewable}_i + \beta_6\,\mathit{ZonePlannedCap}_{t,z} \\ & + \beta_7\,\mathit{Headwinds}_{ff,t} + \beta_8\,\mathit{Headwinds}_{ren,t} \\ & + \beta_9\,\mathit{Turbulence}^{(resid)}_{ff,t} + \beta_{10}\,\mathit{Turbulence}^{(resid)}_{ren,t}\end{aligned} \tag{10}$$

- $\mathit{Nameplate}_i$: Nameplate capacity of generator i. (MW)
- $\mathit{Duration}_{i,t}$: Duration generator i has been in the Planning stage by year t. (years)
- $\mathit{CapPmt}_{t+1,z}$: Capacity (RPM) price in year $t+1$ and zone z. (\$/kW-yr)
- $\mathit{GroupSize}_{i,t}$: Year t total number of generators in group i.
- $\mathit{Renewable}_i$: Indicator variable, equals 1 if generator i is a renewable resource.
- $\mathit{ZonePlannedCap}_{t+1,z}$: Planned year t capacity in zone z.
- $\mathit{Headwinds}_{ff,t}$: Year t value of *Headwinds* for fossil fuel generators, set to zero otherwise.
- $\mathit{Headwinds}_{ren,t}$: Year t value of *Headwinds* for renewable generators, set to zero otherwise.

- $Turbulence_{ff,t}$: Year t value of *Turbulence* for fossil fuel generators, set to zero otherwise.
- $Turbulence_{ren,t}$: Year t value of *Turbulence* for renewable generators, set to zero otherwise.
- $Turbulence^{(resid)}_{ff,t}$: Residualized year t value of *Turbulence* for fossil fuel generators, set to zero otherwise.
- $Turbulence^{(resid)}_{ren,t}$: Residualized year t value of *Turbulence* for renewable generators, set to zero otherwise.

The model clusters standard errors at the state level and sets the base outcome for the multinomial logistic regression as remaining in PLG.

3.2 Full Sample 2008–2023

In the full sample, the variable $\mathbf{Transition}_{i,t+1}$ can take only three values.

- $\mathbf{Transition}_{i,t+1}$: Categorical variable indicating the transition of generator i from the Planning (PLG) stage in year t to year $t+1$. It takes the value of $\{1, 2, 3\}$ if generator i is in the Planning (PLG), Construction (CON), or Indefinitely Postponed/Canceled (IDP) stage in year $t+1$, respectively.

Table 2 presents the full sample coefficients as odds ratios.

3.2.1 Interpretation of Full Sample Results: Table 2

The full-sample odds ratios in Table 2 indicate several consistent patterns across Specifications (1)–(4). Nameplate capacity has a statistically significant but economically modest effect. Larger generators are less likely to transition from PLG to either CON or IDP/CNL, with odds ratios between 0.985–0.989 for CON (all but one $p < 0.05$), and 0.978—0.982 for IDP/CNL (all $p < 0.05$). This suggests that larger units face greater downside risk or higher irreversibility costs, and therefore proceed more cautiously.

Duration is strongly significant in all specifications. A longer time spent in PLG sharply reduces the probability of entering CON (odds 0.62, $p < 0.001$) and raises the probability of IDP/CNL (odds 1.31, $p < 0.01$) . Duration therefore acts as a potent negative signal in the development pipeline.

Capacity payments (CapPmt) increase the likelihood of moving from PLG to CON (odds 1.017, $p = 0.002$) but have no meaningful effect on IDP/CNL transitions (odds 1.006, $p > 0.2$).

Similar to Nameplate, GroupSize reduces the probability of transitioning into CON (odds 0.66–0.69, $p < 0.01$) but has no statistically significant effect on IDP/CNL transitions (odds 0.90–0.94, $p > 0.30$). This indicates that multi-unit proposals tend to move forward more slowly, but once in planning, they are not disproportionately likely to be abandoned.

Renewable projects are significantly less likely than fossil projects to proceed to CON (odds 0.34–0.59, $p < 0.05$).

Table 2 shows that *Headwinds* and *Turbulence* have limited statistical power in the full sample. *Headwinds* for fossil generators have imprecise coefficients (odds 0.69–0.74, p

Table 2: Multinomial Logit Regression (2008–2023 Full Sample, Unregulated Only). Odds ratios with p-values in parentheses. Base outcome = PLG.

	CON				**IDP/CNL**			
Variable	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Nameplate	0.985	0.988	0.989	0.988	0.982	0.978	0.978	0.978
	(0.003)	(0.023)	(0.056)	(0.024)	(0.001)	(0.000)	(0.000)	(0.000)
Duration	0.618	0.621	0.622	0.616	1.304	1.309	1.309	1.308
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
CapPmt	1.017	1.017	1.017	1.018	1.007	1.006	1.007	1.006
	(0.003)	(0.002)	(0.002)	(0.002)	(0.228)	(0.384)	(0.253)	(0.368)
GroupSize	0.664	0.683	0.689	0.680	0.937	0.898	0.913	0.899
	(0.000)	(0.002)	(0.002)	(0.002)	(0.542)	(0.326)	(0.404)	(0.329)
Renewable	0.586	0.365	0.360	0.341	0.222	0.913	0.919	0.907
	(0.037)	(0.027)	(0.042)	(0.022)	(0.000)	(0.884)	(0.902)	(0.880)
ZonePlannedCap	0.988	0.987	0.987	0.987	1.001	1.003	1.003	1.003
	(0.030)	(0.024)	(0.023)	(0.027)	(0.822)	(0.687)	(0.648)	(0.679)
Headwinds (FF)	—	0.737	—	0.694	—	2.023	—	1.995
		(0.404)		(0.339)		(0.056)		(0.078)
Headwinds (Ren)	—	1.438	—	1.339	—	0.380	—	0.349
		(0.383)		(0.503)		(0.057)		(0.036)
Turbulence (FF)	—	—	0.693	0.471	—	—	1.511	0.897
			(0.155)	(0.437)			(0.160)	(0.864)
Turbulence (Ren)	—	—	1.067	0.577	—	—	0.534	0.574
			(0.816)	(0.567)			(0.036)	(0.726)
Constant	2.137	2.397	2.969	2.553	0.271	0.163	0.154	0.164
	(0.120)	(0.072)	(0.032)	(0.057)	(0.035)	(0.008)	(0.012)	(0.009)
Observations	870				870			
Pseudo R^2	0.1034	0.1104	0.1095	0.1113	0.1034	0.1104	0.1095	0.1113
Log Likelihood	-718.24	-712.60	-713.34	-711.94	-718.24	-712.60	-713.34	-711.94
AIC	1464.47	1461.19	1462.68	1467.89	1464.47	1461.19	1462.68	1467.89

Notes: Models (1)–(4) add Headwinds and Turbulence terms in stages, with (4) including both. RRR = exp(
beta); p-values in parentheses. FF = Fossil Fuel, Ren = Renewable. Sample includes unregulated generators only.

= 0.34–0.40 for CON; odds 2.02, p = 0.056 for IDP/CNL) while renewable *Headwinds* exhibit moderate effects (for IDP/CNL, odds 0.35, $p < 0.04$). *Turbulence* coefficients are similarly small and uncertain. The full-sample results therefore suggest that the baseline drivers of transitions are project characteristics (Duration, Nameplate, GroupSize, Renewable) and RPM incentives.

3.3 Subsample 2016–2023

In the subsample, the variable $\textbf{Transition}_{i,t+1}$ can take four values.

- $\textbf{Transition}_{i,t+1}$: Categorical variable indicating the transition of generator i from the Planning (PLG) stage in year t to year $t+1$. It takes the value of {1, 2, 3, 4} if generator i is in the Planning (PLG), Construction (CON), or Indefinitely Postponed (IDP), or Canceled (CNL) stage in year $t+1$, respectively.

Table 3 presents the subsample coefficients as odds ratios.

3.3.1 Interpretation of subsample Results: Table 3

With IDP and CNL separated, several effects become sharper. Nameplate remains negatively associated with transitions into any non-PLG stage, especially cancellations (CNL odds 0.95, $p < 0.001$) . Larger units are thus more likely to linger in PLG.

Duration continues to reduce movement into CON (odds = 0.51, $p < 0.001$) and substantially increases the likelihood of cancellation (CNL odds 1.56–1.65, p = 0.001) but remains insignificant for IDP (odds 1.07, $p > 0.45$) . This confirms that prolonged planning converts primarily into cancellations rather than indefinite postponements once CNL exists as a separate status.

CapPmt again raises the probability of entering CON (odds 1.031, $p < 0.001$) and also slightly raises the probability of IDP (odds 1.035, p = 0.02–0.03) while not affecting cancellations ($p > 0.7$).

GroupSize strongly suppresses transitions into CON (odds 0.38–0.42, $p < 0.005$) and strongly suppresses cancellations (odds 0.40–0.50, $p < 0.001$), but has no effect on IDP ($p > 0.5$).

Across Specifications (1)–(4) in Table 3, the coefficients on Duration, CapPmt, Nameplate, and GroupSize are remarkably stable. Adding Headwinds, Turbulence, or both barely changes either the magnitudes or the signs of these estimates. This robustness strengthens the interpretation that these variables capture fundamental features of the sequential investment process—time in queue, project scale, and capacity-market incentives—rather than artifacts of any particular uncertainty specification.

In specification 4 the coefficients Renewables–for CON, IDP, and CNL–are statistically insignificant.

In contrast with the full sample, *Headwinds* and *Turbulence* have stronger effects in the subsample. They are also very hard to estimate. Because the indices are aggregated across fuels and there are only 23 IDP and 39 CNL transitions, the raw odds ratios in Table 3 (often in the tens or hundreds) should not be taken literally. We therefore focus

Table 3: Multinomial Logit Regression (2016–2023 Subsample, Unregulated Only). Odds ratios with p-values in parentheses. Base outcome = PLG.

	CON				**IDP**				**CNL**			
Variable	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Nameplate	0.986	0.985	0.982	0.985	0.991	0.984	0.987	0.985	0.961	0.953	0.947	0.947
	(0.050)	(0.059)	(0.027)	(0.087)	(0.352)	(0.128)	(0.202)	(0.144)	(0.000)	(0.000)	(0.000)	(0.000)
Duration	0.505	0.510	0.502	0.511	1.103	1.087	1.096	1.068	1.565	1.646	1.617	1.612
	(0.000)	(0.000)	(0.000)	(0.000)	(0.460)	(0.596)	(0.499)	(0.680)	(0.001)	(0.001)	(0.001)	(0.002)
CapPmt	1.025	1.025	1.024	1.031	1.030	1.018	1.028	1.035	1.004	0.998	1.006	1.003
	(0.000)	(0.000)	(0.000)	(0.000)	(0.016)	(0.218)	(0.030)	(0.006)	(0.725)	(0.877)	(0.671)	(0.826)
GroupSize	0.405	0.397	0.375	0.418	1.114	0.976	1.047	1.036	0.492	0.414	0.415	0.399
	(0.002)	(0.002)	(0.002)	(0.004)	(0.563)	(0.910)	(0.820)	(0.863)	(0.000)	(0.000)	(0.000)	(0.000)
Renewable	0.275	0.299	0.881	0.482	0.225	54.786	1.509	14.637	0.065	9.121	30.286	27.335
	(0.026)	(0.178)	(0.878)	(0.402)	(0.013)	(0.346)	(0.792)	(0.425)	(0.000)	(0.236)	(0.144)	(0.257)
ZonePlannedCap	0.989	0.989	0.989	0.988	1.005	1.004	1.005	1.006	0.998	0.998	0.999	0.999
	(0.099)	(0.089)	(0.095)	(0.074)	(0.713)	(0.767)	(0.742)	(0.652)	(0.877)	(0.885)	(0.935)	(0.940)
Headwinds (FF)	—	1.672	—	2.713	—	532.354	—	98.607	—	17.127	—	49.523
		(0.444)		(0.167)		(0.006)		(0.007)		(0.046)		(0.095)
Headwinds (Ren)	—	1.906	—	1.295	—	66.973	—	10.241	—	0.226	—	0.202
		(0.245)		(0.662)		(0.092)		(0.375)		(0.034)		(0.033)
Turbulence (FF)	—	—	2.414	18.151	—	—	4.298	0.647	—	—	13.893	12.667
			(0.066)	(0.037)			(0.020)	(0.892)			(0.043)	(0.187)
Turbulence (Ren)	—	—	1.205	0.109	—	—	1.663	0.002	—	—	0.352	0.333
			(0.597)	(0.119)			(0.471)	(0.056)			(0.012)	(0.715)
Constant	7.715	4.189	2.159	2.109	0.033	0.000	0.003	0.000	0.896	0.032	0.009	0.010
	(0.033)	(0.143)	(0.383)	(0.453)	(0.003)	(0.001)	(0.000)	(0.000)	(0.929)	(0.072)	(0.025)	(0.082)
Observations	562				562				562			
Pseudo R^2	0.1767	0.1998	0.1918	0.2109	0.1767	0.1998	0.1918	0.2109	0.1767	0.1998	0.1918	0.2109
Log Likelihood	-452.98	-440.25	-444.66	-434.15	-452.98	-440.25	-444.66	-434.15	-452.98	-440.25	-444.66	-434.15
AIC	947.96	934.51	943.33	934.29	947.96	934.51	943.33	934.29	947.96	934.51	943.33	934.29

Notes: Headwinds and Turbulence split by FF = Fossil Fuel and Ren = Renewable. Sample includes only unregulated generators in the 2016–2023 subsample.

on the signs and on the marginal effects in Section 4 rather than on the magnitudes of the coefficients.

We next turn to the marginal effects, which visualize these relationships and reveal non-linearities that the odds ratios alone cannot show.

4 Marginal Effects of Duration, Capacity Prices, Group Size, and Uncertainty

All marginal effects are based upon Specification 4 in equation (10). We focus on the subsample (2016–2023) for clarity.

4.1 Marginal Effects of Duration

Figure 3 show that the marginal effects of *Duration* change sign depending on how long the generator has been in the Planning (PLG) stage. For generators that have recently entered PLG, a small increase in *Duration* increases the probability of remaining in PLG and decreases the probability of moving to Construction (CON).

After roughly 3–4 years in PLG, an additional year of *Duration* makes it less likely for the generator to remain in PLG and more likely to exit planning (PLG) via cancellation (CNL).

4.2 Marginal Effects of Capacity (RPM) Prices

Figure 4 shows the marginal effects of capacity prices on transitions in the 2016–2023 subsample. Higher capacity prices reduce the probability of remaining in PLG and increase the probability of entering CON across the full range of observed prices, with essentially no effect on cancellations (CNL). This result supports the interpretation that RPM capacity payments provide a forward-looking investment incentive rather than simply flushing projects out of the queue.

4.3 Marginal Effects of Group Size

Figure 5 displays the marginal effects of GroupSize. Larger groups are less likely to transition to CON and less likely to be canceled, with little change in IDP.

Figure 6 displays the marginal effects of Nameplate capacity.

4.4 Marginal Effects of Uncertainty

Figures 7 and 8 display the marginal effects of *Turbulence*. For fossil generators, higher *Turbulence* reduces the probability of remaining in PLG and raises the probability of moving to CON. Renewable generators' response to higher *Turbulence* show the opposite–increased values of *Turbulence* are associated with remaining higher probabilities of remaining in PLG and lower probabilities moving to CON. However, the renewable marginal effects are small in magnitude and the confidence bands are wide, so we interpret these patterns as suggestive rather than definitive.

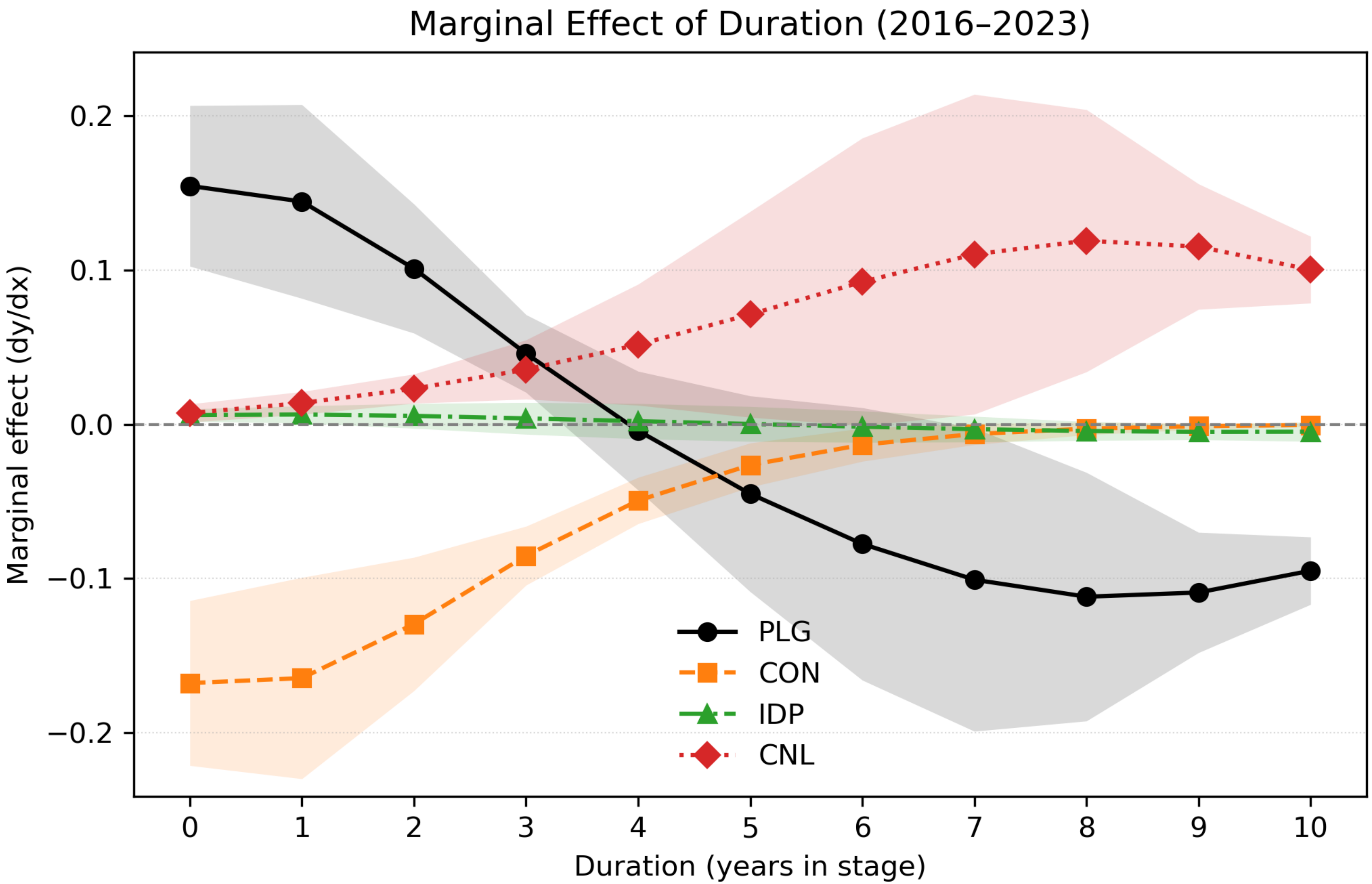

Figure 3: The plot shows the marginal effect of an increase in Duration. Subsample (2016-2023). PLG = Planning; CON = Construction; IDP = Indefinitely Postponed; CNL = Canceled.

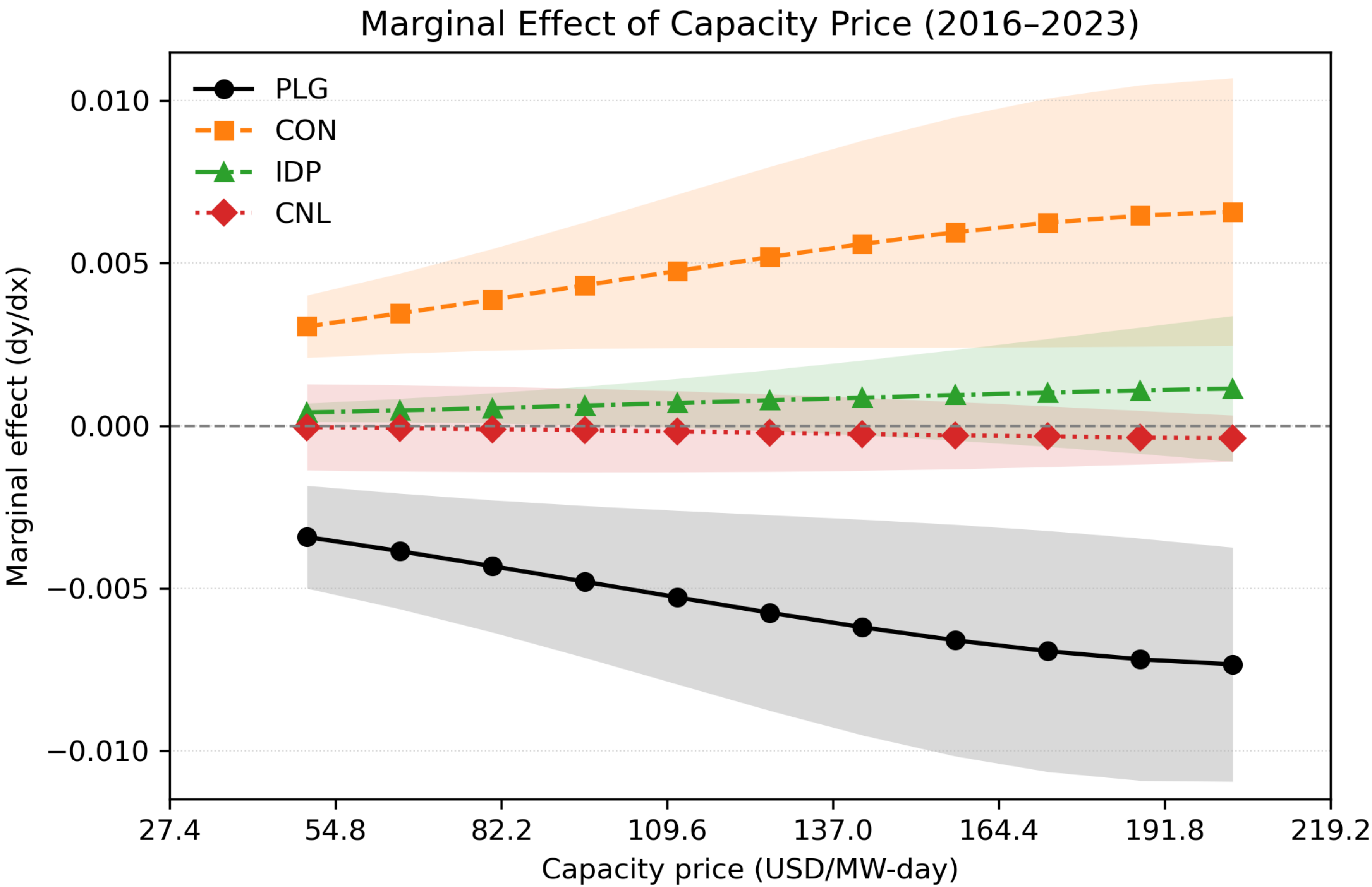

Figure 4: The plot shows the marginal effect of an increase in CapPmt. Subsample (2016-2023). PLG = Planning; CON = Construction; IDP = Indefinitely Postponed; CNL = Canceled.

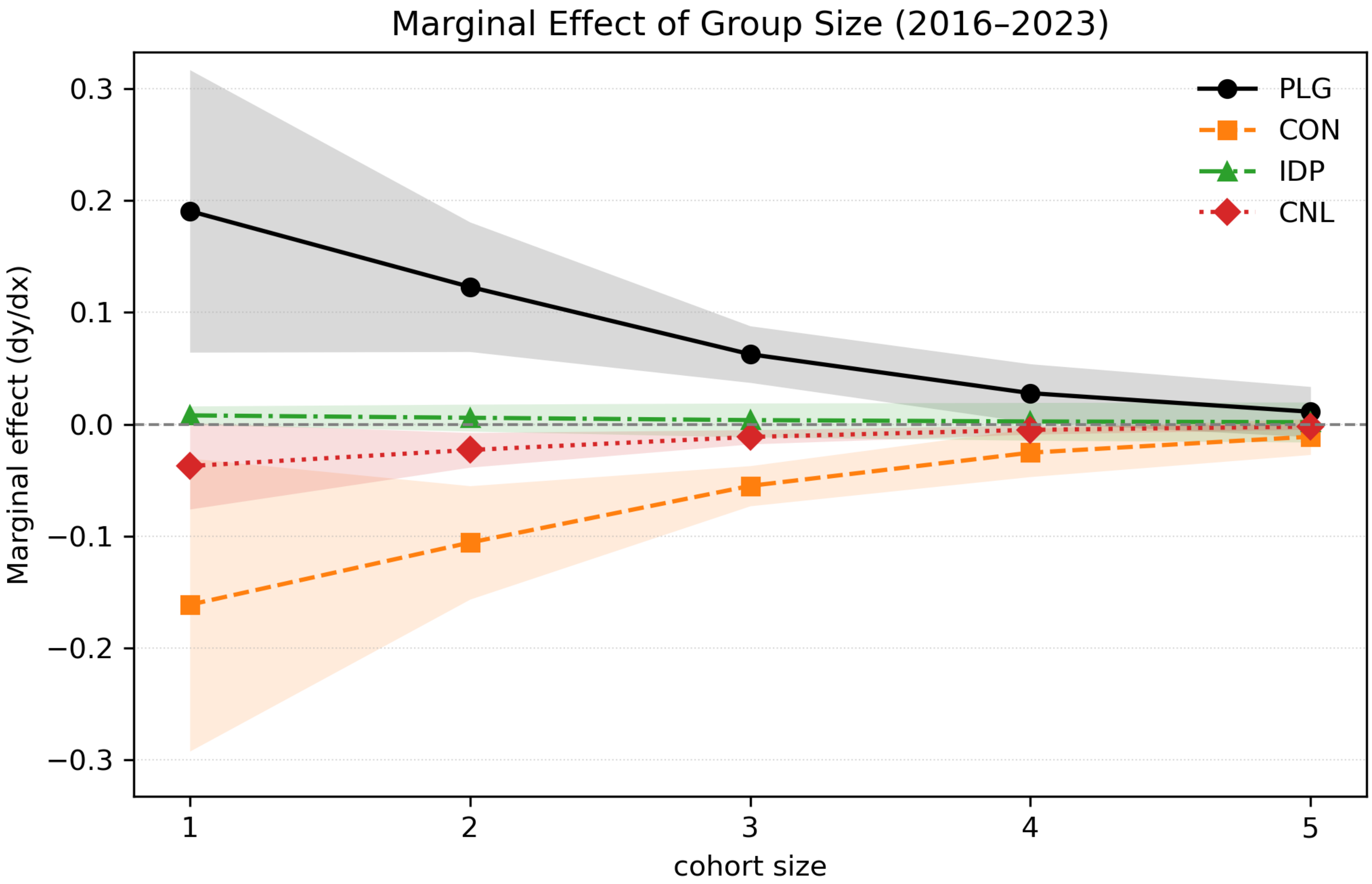

Figure 5: The plot shows the marginal effect of an increase in GroupSize. Subsample (2016-2023). PLG = Planning; CON = Construction; IDP = Indefinitely Postponed; CNL = Canceled.

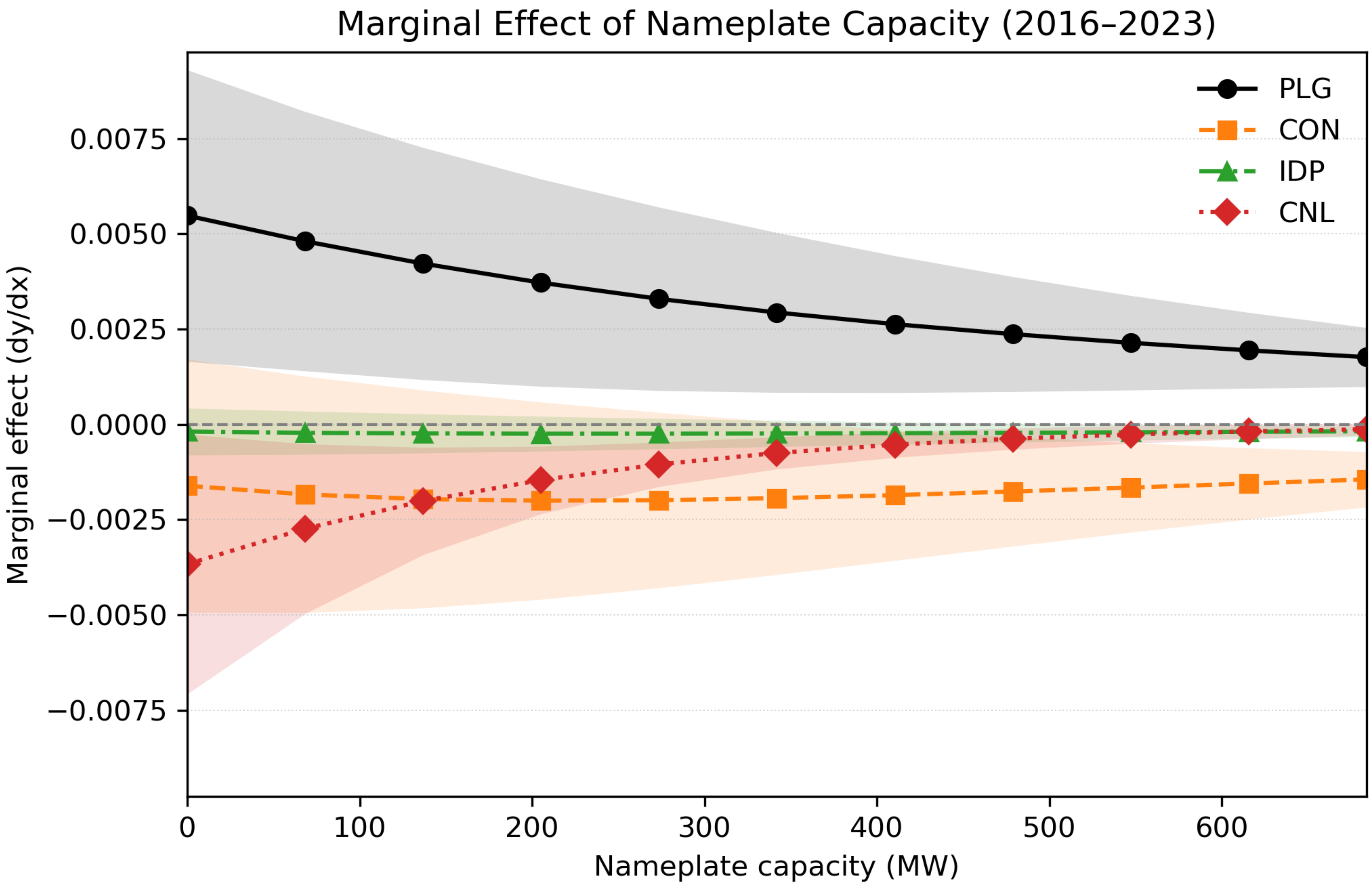

Figure 6: The plot shows the marginal effect of an increase in Nameplate capacity. Subsample (2016-2023). PLG = Planning; CON = Construction; IDP = Indefinitely Postponed; CNL = Canceled.

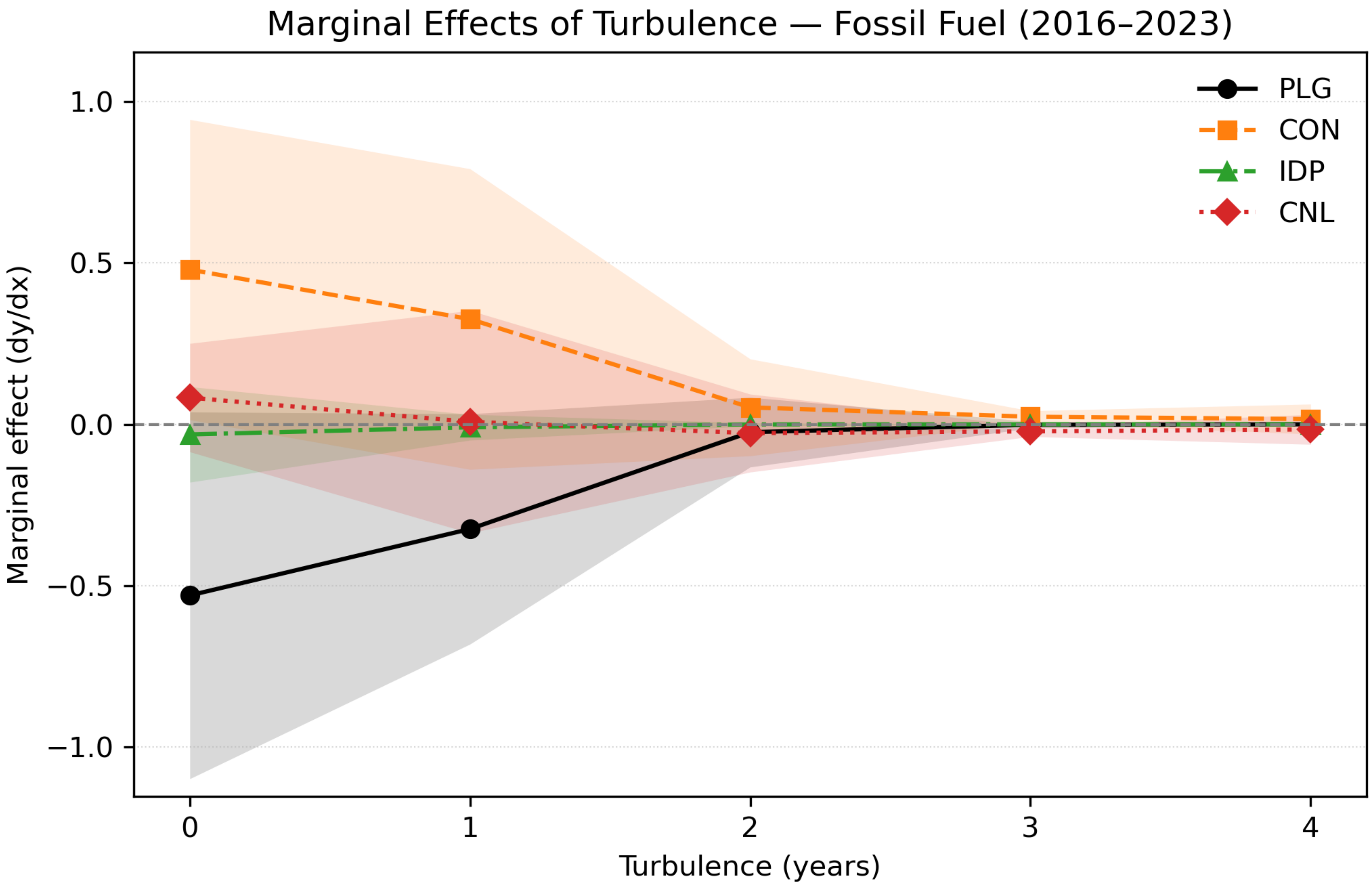

Figure 7: The plot shows the marginal effect of an increase in Turbulence for fossil fuel generators. The sample period is 2016–2023. PLG = Planning; CON = Construction; IDP = Indefinitely Postponed; CNL = Canceled.

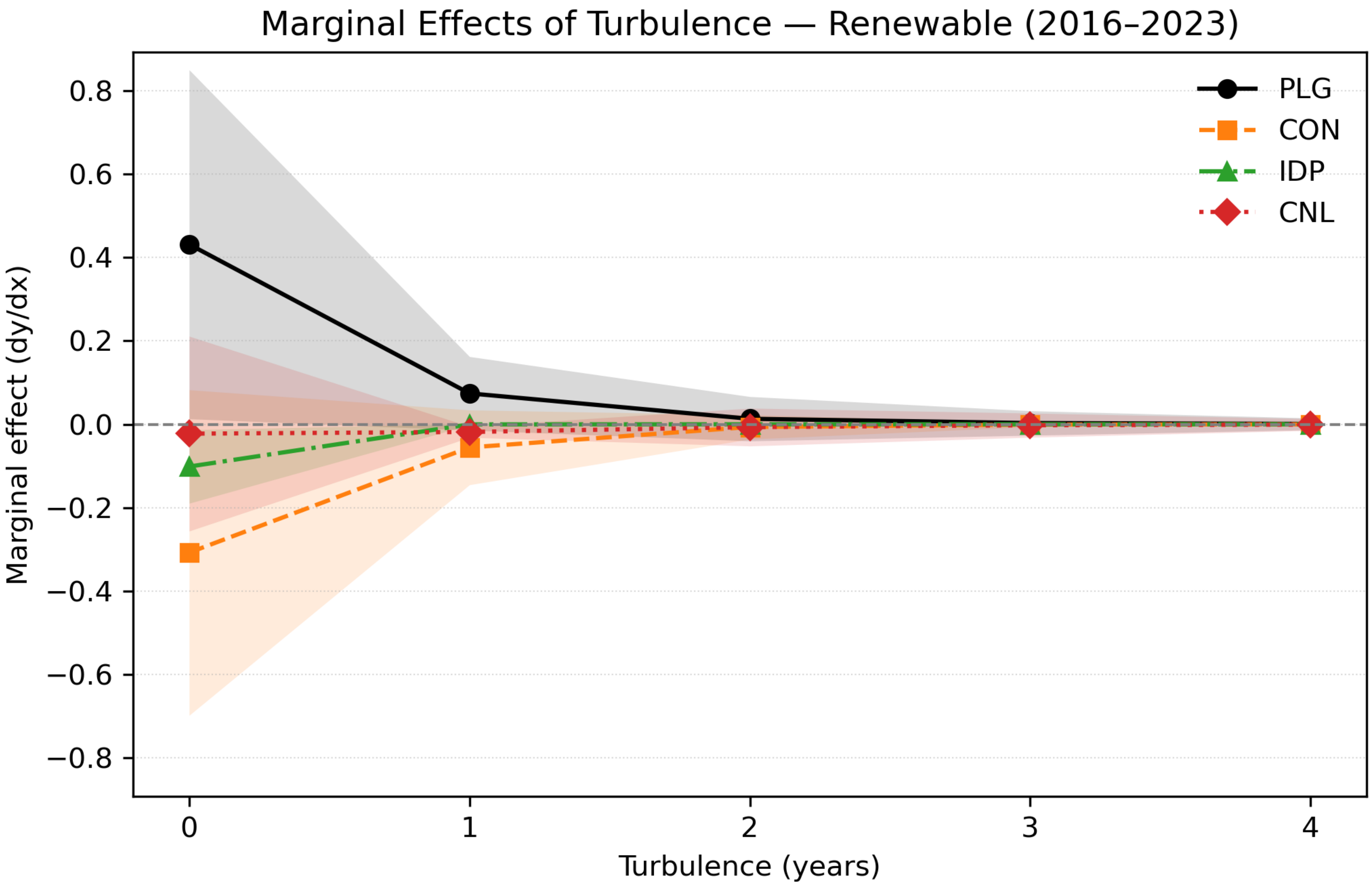

Figure 8: The plot shows the marginal effect of an increase in Turbulence for renewable generators. The sample period is 2016–2023. PLG = Planning; CON = Construction; IDP = Indefinitely Postponed; CNL = Canceled.

5 Conclusion

We study sequential investment in proposed electric power generators in PJM and document how projects move from Planning into Construction, indefinite postponement, and cancellation. Duration in the Planning stage, project scale, and capacity prices emerge as stable, economically meaningful drivers of these transitions. Projects that linger in Planning become increasingly likely to cancel; larger and multi-unit projects move more slowly and are harder to abandon; and higher RPM capacity prices consistently shift generators from Planning into Construction, without increasing Cancellations.

Our Headwinds and Turbulence indices, derived from project-level revisions in expected in-service dates and constructed from non-PJM data via Bayesian partial pooling, provide technology-specific measures of drag and uncertainty in the planning environment. Although their coefficients are large and imprecise in the 2016–2023 subsample, the marginal effects suggest that higher Turbulence pushes fossil-fuel projects out of Planning and into Construction, while renewable responses are muted and tilt toward remaining in Planning. Taken together, these patterns show that simple statements such as "uncertainty delays investment" are incomplete for modern electricity markets.

An avenue for further research would be to extend the study by expanding the sample to be nationwide. We are actively working on this extension. For now, we caution that our results are specific to PJM and may not generalize to other regions.

6 Acknowledgments

This project grew out of earlier work by Kristoffer Ingebrigtsen and Jonas Aase Kaldahl. Their efforts were a major impetus. We thank Erik Haugom for valuable feedback. We thank Ella Santillo for excellent research assistance. All remaining errors are the sole responsibility of the authors.

We acknowledge support from the Research Council of Norway through projects 228811 and 268093, and recognize the Norwegian research centre CenSES, Centre for Sustainable Energy Studies (RCN grant 209697).

References

Baker, S. R., Bloom, N. and Davis, S. J. (2016), 'Measuring economic policy uncertainty', *The Quarterly Journal of Economics* **131**(4), 1593–1636.

Bar-Ilan, A. and Strange, W. C. (1998), 'A model of sequential investment', *Journal of Economic Dynamics and Control* **22**(3), 437–463.

Bulan, L., Mayer, C. and Somerville, C. T. (2009), 'Irreversible investment, real options, and competition: Evidence from real estate development', *Journal of Urban Economics* **65**(3), 237–251.

Bulan, L. T. (2005), 'Real options, irreversible investment and firm uncertainty: New evidence from U.S. firms', *Review of Financial Economics* **14**(3–4), 255–279.

Çam, E., Hinkel, N. and Schönfisch, M. (2022), A stochastic discrete choice dynamic programming model of power plant operations and retirement, Technical report, EWI Working Paper.

Chronopoulos, M., Hagspiel, V. and Fleten, S.-E. (2016), 'Stepwise green investment under policy uncertainty.', *Energy Journal* **37**(4).

Chronopoulos, M., Hagspiel, V. and Fleten, S.-E. (2017), 'Stepwise investment and capacity sizing under uncertainty', *OR spectrum* **39**(2), 447–472.

Cook, J. A. and Lin Lawell, C.-Y. C. (2020), 'Wind turbine shutdowns and upgrades in denmark: Timing decisions and the impact of government policy', *The Energy Journal* **41**(3), 81–118.

Dixit, A. and Pindyck, R. S. (1994), *Investment Under Uncertainty*, Princeton University Press.

Elliott, J. T. (2022), Investment, emissions, and reliability in electricity markets, Technical report, Working paper.

Fabrizio, K. R. (2013), 'The effect of regulatory uncertainty on investment: Evidence from renewable energy generation', *The Journal of Law, Economics, and Organization* **29**(4), 765–798.

Favero, C. A., Hashem Pesaran, M. and Sharma, S. (1994), 'A duration model of irreversible oil investment: Theory and empirical evidence', *Journal of Applied Econometrics* **9**(S1), S95–S112.

Fleten, S.-E., Fram, B. P. and Ullrich, C. J. (2024), 'The reliability pricing model and coal-fired generators in pjm', *Energy Economics* **134**, 107543.

Fleten, S.-E., Haugom, E., Pichler, A. and Ullrich, C. J. (2020), 'Structural estimation of switching costs for peaking power plants', *European Journal of Operational Research* **285**(1), 23–33.

Fleten, S.-E., Haugom, E. and Ullrich, C. (2017), 'The real options to shutdown, startup, and abandon: U.S. electricity industry evidence', *Energy Economics* **63**, 1–12.

Fleten, S.-E., Linnerud, K., Molnár, P. and Nygaard, M. T. (2016), 'Green electricity investment timing in practice: Real options or net present value?', *Energy* **116**, 498–506.

Fleten, S.-E., Maribu, K. M. and Wangensteen, I. (2007), 'Optimal investment strategies in decentralized renewable power generation under uncertainty', *Energy* **32**(5), 803–815.

Gollier, C., Proult, D., Thais, F. and Walgenwitcz, G. (2005), 'Choice of nuclear power investment under price uncertainty: Valuing modularity', *Energy Economics* **27**(4), 667–685.

Huchzermeier, A. and Loch, C. H. (2001), 'Project management under risk: Using the real options approach to evaluate flexibility in R&D', *Management Science* **47**(1), 85–101.

Ishii, J. and Yan, J. (2011), 'Investment under regulatory uncertainty: U.S. electricity generation investment 1996–2000', *UCEI CSEM Working paper #127* .

Kellogg, R. (2014), 'The effect of uncertainty on investment: Evidence from Texas oil drilling', *American Economic Review* **104**(6), 1698–1734.

Linnerud, K., Andersson, A. M. and Fleten, S.-E. (2014), 'Investment timing under uncertain renewable energy policy: an empirical study of small hydropower projects', *Energy* **78**, 154–164.

Majd, S. and Pindyck, R. (1987), 'Time to build, option value, and investment decisions', *Journal of Financial Economics* **40**, 7–27.

Marmer, V. and Slade, M. E. (2018), 'Investment and uncertainty with time to build: Evidence from entry into U.S. copper mining', *Journal of Economic Dynamics and Control* **95**, 233–254.

McDonald, R. and Siegel, D. (1986), 'The value of waiting to invest', *Quarterly Journal of Economics* **101**, 707–728.

Moel, A. and Tufano, P. (2002), 'When are real options exercised? An empirical study of mine closing', *Review of Financial Studies* **1**(15), 35–64.

Näsäkkälä, E. and Fleten, S.-E. (2005), 'Flexibility and technology choice in gas fired power plant investments', *Review of Financial Economics* **14**(2), 371–393.

Paddock, J., Siegel, D. and Smith, J. (1988), 'Option valuations of claims on real assets: The case of offshore petroleum leases', *The Quarterly Journal of Economics* **103**(3), 88–100.

Quigg, L. (1993), 'Empirical testing of real option-pricing models', *The Journal of Finance* **47**(2), 621–640.

Roberts, K. and Weitzman, M. L. (1981), 'Funding criteria for research, development, and exploration projects', *Econometrica* **49**(5), 1261–1288.

Siddiqui, A. and Fleten, S.-E. (2010), 'How to proceed with competing alternative energy technologies: A real options analysis', *Energy Economics* **32**(4), 817–830.

Somerville, C. T. (2001), 'Permits, starts, and completions: structural relationships versus real options', *Real Estate Economics* **29**(1), 161–190.

Walls, W., Rusco, F. W. and Ludwigson, J. (2007), 'Power plant investment in restructured markets', *Energy* **32**(7), 1403–1413.

A Data Cleaning Process

In this appendix, we describe the key steps undertaken to clean the dataset used in our analysis of sequential investment in electric power generators. These steps were necessary to address inconsistencies and ensure the quality of the data for regression analysis.

The raw data come directly from EIA form 860. We retain the variables in Table 4 from EIA 860. The **variables in bold** are relevant for the cleaning process.

Table 4: EIA Form 860 Variables and Their Definitions

Variable	**Definition**
year	EIA 860 reporting year
pcode	Plant Code (unique identifier for the plant)
gcode	Generator Code (unique, within each `pcode`, identifier for the generator)
iso	Independent System Operator region
regst	Regulatory Status (regulated/unregulated)
state	State abbreviation where the plant is located
zip	Five digit ZIP code of the plant
pm	Prime Mover code (see Table **??**)
STATUS	Status of the generator
nplate	Nameplate capacity of the generator (MW)
efmn	Effective Month: Original projected in-service month
efyr	Effective Year: Original projected in-service year
cumn	Current Month: Most recent projected in-service month
cuyr	Current Year: Most recent projected in-service year
nrg1	Primary energy source (fuel) code
nrg2	Secondary energy source (fuel) code

We address several data issues, detailed below, before conducting the analysis.

A.1 Status Code List

Below is a complete list of STATUS codes in EIA860.

- `P` (Proposed)
- `L` (Permitted)
- `T` (Construction)
- `U` (Under Construction)

- `V` (Verification)
- `TS` (Testing)
- `IP` (In Progress)
- `OT` (Other)
- `OP` (Operational)
- `BU` (Backup)
- `SB` (Standby)
- `OA` (Out of Action)
- `OS` (Out of Service)
- `RE` (Retired)
- `ZZ` (Scrapped)
- `CN` (Canceled)

A.2 Prime Mover (PM) Cleaning

Prime mover codes, which represent the technology used by the generator, required careful handling due to missing or inconsistent entries. We apply the procedure for all generators in the database, even if they were already existing (and therefore not in the Planning or Construction stages) at any point in the sample.

- We implemented a function that checks for missing PM values and fills them if non-missing PMs are consistent across time for a given generator. After applying this cleaning procedure, 1,948 generators had missing PMs that were successfully filled with consistent values.
- We logged and reviewed each case where a PM was changed to ensure that the process was applied correctly. We reviewed all such cases to make the determination whether the PM change was legitimate or not. In the following we provide some common examples.
  - In the EIA 860, GT is used for simple cycle combustion turbines andx CT is used for combustion turbines which are a part for a Combined Cycle. It is common in the industry to use these acronyms interchangeably. Not surprisingly, these two PM codes get mixed up often in the data.
  - Sometimes an existing simple cycle combustion turbine (GT) is converted to combined cycle operation (converted to a CT) with the addition of a steam turbine[13].

A.3 Gcode Changes

We also identified and corrected potential changes in generator codes (gcodes) over time. Using a tolerance threshold, we flagged gcodes with similar capacity (`nplate`) and con-

[13] A heat recovery steam generator.

sistent status code sequences as candidates for consolidation. A total of 194 potential gcode changes were identified and reviewed, with the feedback process allowing us to apply these changes as appropriate.

A.4 Combining Combined Cycle (CC) Components

For generators utilizing combined cycle technology, it was necessary to combine parts of the cycle (e.g., gas turbines (CT), steam turbines (CA), and combined cycle components (CS)) into a single entry. We used status code sequences and capacity data to identify similar sequences and consolidated them into a single generator entry. This step was crucial for ensuring that combined cycle generators were treated consistently in the analysis.

A.5 Proposed Generators Filtering

We filtered the dataset to retain only generators that were proposed at some point in time. We defined a generator as "proposed" if any of its status codes fell within a specific range (e.g., `P`, `L`, `T`, `U`, `V`, etc.).

A.6 Iterative Process

The data cleaning process was designed to be iterative. For example, changes in Prime Mover (PM) values could influence gcode changes and the combining of Combined Cycle components. Each step in the process was carefully logged and reviewed before moving to the next, ensuring that feedback from one step could be incorporated into earlier steps as needed.

This cleaning process is ongoing, and future refinements will likely be made as the analysis progresses.

B Bayesian Partial Pooling

Standardization, Trimming, and Weighting

Before fitting the Bayesian hierarchical model, we standardize the project-level delays within each technology group. For each technology k, we compute the raw mean m_k and standard deviation s_k of `DiffYear`, and transform the observations

$$y_{i,k,t} = \frac{\texttt{DiffYear}_{i,t} - m_k}{s_k}.$$

This ensures that groups with different units (e.g., wind vs. coal) enter the model on a comparable scale. We then trim the standardized observations as $|y_{i,k,t}| \leq z_{\max}$, removing extreme outliers or data errors. The choice of $z_{\max}$ is configurable and set to 3 for the indices used in this paper.

Hierarchical Model for Technology–Year Means

For each technology $k \in \{1, \dots, 2\}$ and year $t \in \{1, \dots, T\}$ the model posits a latent mean $\mu_{k,t}$ and standard deviation $\sigma_{k,t}$ for the standardized delays:

$$y_{i,k,t} \sim \mathcal{N}\big(\mu_{k,t},\, \sigma_{k,t}^2\big).$$

The means are decomposed into a global intercept, a technology effect, a common year effect, and a technology–year deviation:

$$\mu_{k,t} = \alpha + u_{\mu,k} + g_{\mu,t} + v_{\mu,k,t}, \tag{11}$$

where

- α is a global intercept for the standardized delays;
- $u_{\mu,k}$ is a technology–specific random effect;
- $g_{\mu,t}$ is a year effect common to all technologies; and
- $v_{\mu,k,t}$ is a technology–year deviation.

The random effects are given Normal priors with unknown scales

$$u_{\mu,k} \sim \mathcal{N}\big(0, \sigma_{u,\mu}^2\big), \qquad g_{\mu,t} \sim \mathcal{N}\big(0, \sigma_{g,\mu}^2\big), \qquad v_{\mu,k,t} \sim \mathcal{N}\big(0, \sigma_{\mu,k}^2\big). \tag{12}$$

To improve sampling, the Stan implementation uses a non–centered parameterization, writing each effect as a standard Normal draw multiplied by its scale (e.g. $u_{\mu,k} = \sigma_{u,\mu}\, \tilde{u}_{\mu,k}$ with $\tilde{u}_{\mu,k} \sim \mathcal{N}(0,1)$). The year effects $g_{\mu,t}$ are centered across years in the transformed–parameters block so that α represents the overall mean level of the standardized delays across all technologies and years.

Hierarchical Model for Technology–Year Volatility

The dispersion of delays within each technology–year cell is modeled through a log–standard deviation equation. Let

$$\ell_{k,t} \equiv \log \sigma_{k,t}.$$

We specify

$$\ell_{k,t} = \beta_0 + u_{\sigma,k} + g_{\sigma,t} + v_{\sigma,k,t}, \tag{13}$$

with

- β_0 a global intercept for log–volatility;
- $u_{\sigma,k}$ a technology–specific random effect;
- $g_{\sigma,t}$ a year effect common to all technologies; and
- $v_{\sigma,k,t}$ a technology–year deviation.

The corresponding priors mirror those for the mean equation:

$$u_{\sigma,k} \sim \mathcal{N}\big(0, \sigma_{u,\sigma}^2\big), \qquad g_{\sigma,t} \sim \mathcal{N}\big(0, \sigma_{g,\sigma}^2\big), \qquad v_{\sigma,k,t} \sim \mathcal{N}\big(0, \sigma_{\sigma,k}^2\big). \tag{14}$$

Again we use a non–centered parameterization in Stan (e.g. $u_{\sigma,k} = \sigma_{u,\sigma}\,\tilde{u}_{\sigma,k}$ with $\tilde{u}_{\sigma,k} \sim \mathcal{N}(0,1)$), and the year effects $g_{\sigma,t}$ are centered across t so that β_0 corresponds to the overall average log–volatility.

The group-, year-, and cell-level scales are modeled on the log scale. Writing

$$\sigma_{u,\mu} = \exp(\tau_{u,\mu}), \quad \sigma_{u,\sigma} = \exp(\tau_{u,\sigma}), \quad \sigma_{g,\mu} = \exp(\tau_{g,\mu}),$$

$$\sigma_{g,\sigma} = \exp(\tau_{g,\sigma}), \quad \sigma_{\mu,k} = \exp(\tau_{\mu,k}), \quad \sigma_{\sigma,k} = \exp(\tau_{\sigma,k}),$$

we place weakly informative Normal priors

$$\tau_{u,\mu},\ \tau_{u,\sigma},\ \tau_{g,\mu},\ \tau_{g,\sigma} \sim \mathcal{N}\big(\log 0.10,\ 0.25^2\big), \tag{15}$$

$$\tau_{\mu,k},\ \tau_{\sigma,k} \sim \mathcal{N}\big(\log 0.10,\ 0.25^2\big) \qquad \text{independently for each } k. \tag{16}$$

Because $y_{i,k,t}$ has been standardized within each technology, the intercepts are given tight priors,

$$\alpha \sim \mathcal{N}(0, 0.20^2), \qquad \beta_0 \sim \mathcal{N}(0, 0.20^2).$$

Weighted Likelihood

Each observation is given weight w_i in the likelihood, so that the log posterior kernel includes the term

$$\log L = \sum_i w_i \log \phi\big(y_{i,k,t};\ \mu_{k,t}, \sigma_{k,t}\big),$$

where $\phi(\cdot;\mu,\sigma)$ denotes the Normal density. The weights w_i match those used in the construction of the raw indices (e.g. equal weights across blocks or weights proportional to block nameplate capacity).

Recovery of Pooled Headwinds and Turbulence

After fitting the model, we recover posterior means of $\mu_{k,t}$ and $\ell_{k,t}$ (or of $\sigma_{k,t}$ and transform back to the original units of `DiffYear` using the same group-specific standardization constants m_k and s_k that were used to create the standardized observations:

$$\mathit{Headwinds}_{k,t}^{\text{pooled}} = m_k + s_k \cdot \mu_{k,t}, \tag{17}$$

$$\mathit{Turbulence}_{k,t}^{\text{pooled}} = s_k \cdot \exp(\ell_{k,t}). \tag{18}$$

Advantages of the Hierarchical Approach

The Bayesian partial pooling approach provides several benefits relative to raw sample statistics:

1. *Stabilization in thin cells.* Year–technology cells with few or single observations borrow strength from richer years, avoiding extreme values.
2. *Technology-specific flexibility.* The hierarchical model allows both fossil fuel and renewables to have its own time-series pattern with its own pooling scale, avoiding both over-pooling and under-pooling.

Posterior means, posterior standard deviations, and (if desired) credible intervals are available for each technology–year cell, although the main regressions use posterior means as point estimates.

The practical effect of these advantages can be seen in the raw vs. pooled plots below, where the hierarchical model substantially reduces noise while preserving technology-specific trends.

For transparency and reproducibility, the replication files include both raw (unpooled) and pooled versions of *Headwinds* and *Turbulence*. The regression specifications below use the pooled versions.

Raw vs. Pooled Indices and LOO-State Cross-Fit Validation

To illustrate the effect of the Bayesian pooling procedure, Appendix C reports both the raw indices and the pooled indices for each technology group $k \in \{1, 2\}$. For each technology we report four series:

1. **Raw–nationwide:** the unpooled sample mean and sample standard deviation of `DiffYear` using all 51 (including DC) states;
2. **Raw–PJM states omitted:** the same raw indices recomputed after omitting the 14 states which make up the regression sample below.
3. **Pooled–nationwide:** the Bayesian pooled indices using all states in the hierarchical model; and
4. **Pooled–PJM states omitted:** the pooled indices obtained from the cross-fitted Stan model that excludes the regression states.

These figures make two properties of our approach transparent. First, the raw indices (*Raw–nationwide*) can be noisy, especially in early years for technologies with few observations (i.e, renewable). The pooled series (*Pooled–nationwide*) smooth these irregularities while preserving genuine temporal structure. Second, the LOO cross-fit curves (*Raw–PJM states omitted* and *Pooled–PJM states omitted*) provide an external validation of the indices. The hierarchical structure is not overfitting to the regression states. Figures 9 through 12 plot the indices.

Interpreting the Raw and Pooled Index Plots. This appendix provides a visual comparison of the raw indices and the Bayesian partially pooled indices. For each group we show four series—Raw–nationwide, Raw–PJM states omitted, Pooled–nationwide, and Pooled–PJM states omitted—following the definitions in Section 2.4. In all the plots, *nationwide* refers to all 51 states (including DC) and *PJM states omitted* refers to the indices calculated from the 37 states not in our regression samples.

Event Alignment Figures 9 through 12 show that the *Headwinds* and *Turbulence* indices capture the broad timing of several major policy and market shocks listed in Table 5, though not every event generates a sharp spike.

For fossil fuels, *Headwinds* rise steadily from 2001 and reach their maximum in 2005, a period that coincides with the onset of widespread coal project cancellations and growing climate and regulatory pressure. Fossil *Headwinds* then remain elevated through the mid-2000s and only gradually decline after 2010.

For fossil *Turbulence*, the index climbs from the early 2000s into a prolonged plateau between roughly 2004 and 2011, encompassing the 2008–2009 financial crisis, the RGGI launch, and the implementation of CSAPR/MATS. Rather than isolated spikes at each event year, the fossil *Turbulence* series reflects a sustained period of high cross-sectional dispersion in project delays during this era, followed by a second, smaller bump around 2020 that lines up with the COVID-19 disruptions. The 2022 fuel-price shock occurs against this already elevated post-COVID backdrop.

On the renewable side, *Turbulence* is relatively moderate in the early 2000s and then shows a clear local surge around the 2012 wind PTC lapse. *Turbulence* peaks again in 2017–2018, coinciding with the Section 201 solar trade case and subsequent solar tariffs, before easing somewhat thereafter. These two episodes line up well with the policy-driven boom–bust cycles for wind and utility-scale solar. By contrast, the 2020 COVID shock and the 2022 fuel-price and policy changes do not produce sharp spikes in the aggregated renewable *Turbulence* index; they occur during a period of already elevated but relatively stable uncertainty.

Overall, the indices do not imply a one-to-one mapping from each event in Table 5 to a distinct peak, but they do capture the main episodes when planning delays and uncertainty were unusually high.

Table 5: Events annotated in Headwinds and Turbulence plots. These events are major policy and market shocks that are likely to influence planning delays and uncertainty; some correspond to visible changes in the indices in Figures 9 through 12, while others occur during broader periods of elevated or declining Headwinds and Turbulence.

Year	**Event**	**Description**
2005	Coal cancellations	Beginning of widespread coal cancellations driven by environmental regulation, grassroots opposition, and deteriorating project economics.
2007	Mass v. EPA	Supreme Court ruling establishing EPA authority to regulate greenhouse gases under the Clean Air Act.
2008	Financial crisis	Global recession and credit shock, sharply reducing electricity demand and delaying capital-intensive infrastructure.
2009	RGGI launch	Regional Greenhouse Gas Initiative takes effect, establishing a carbon price in the Northeast.
2011	CSAPR/MATS	EPA finalizes two major pollution rules: Cross-State Air Pollution Rule and Mercury and Air Toxics Standards.
2012	Wind PTC lapse	Expiration of wind production tax credit; causes major pipeline contraction and investment pause.
2013	Wind PTC restored	Reinstatement of the PTC via fiscal cliff bill; activity resumes but with uncertainty.
2014–2015	Clean Power Plan and PJM reforms	Announcement of CPP targets + PJM's stricter capacity accreditation rules following the polar vortex.
2016	PTC phaseout begins	Legislative deal begins gradual phase-down of wind credits, providing long-term investment certainty.
2017	Solar tariff probe	Section 201 trade complaint triggers procurement delays and risk repricing for solar developers.
2018	Solar tariffs imposed	President Trump imposes 30% import tariffs on solar panels under Section 201.
2020	COVID-19 pandemic	Global shutdowns disrupt supply chains, permitting, and construction across all fuel types.
2022	Russia–Ukraine crisis	International gas supply shock triggers extreme price volatility; alters fuel risk calculus.
2022.5	Inflation Reduction Act	Historic climate bill delivers 10-year tax credit certainty; reshapes expectations for renewable buildout.
2021–2022	PJM queue freeze	PJM halts interconnection intake due to study backlog; delays hundreds of GW.
2023	PJM reforms restart	Queue restarts under new cluster-based process, but backlog remains substantial.

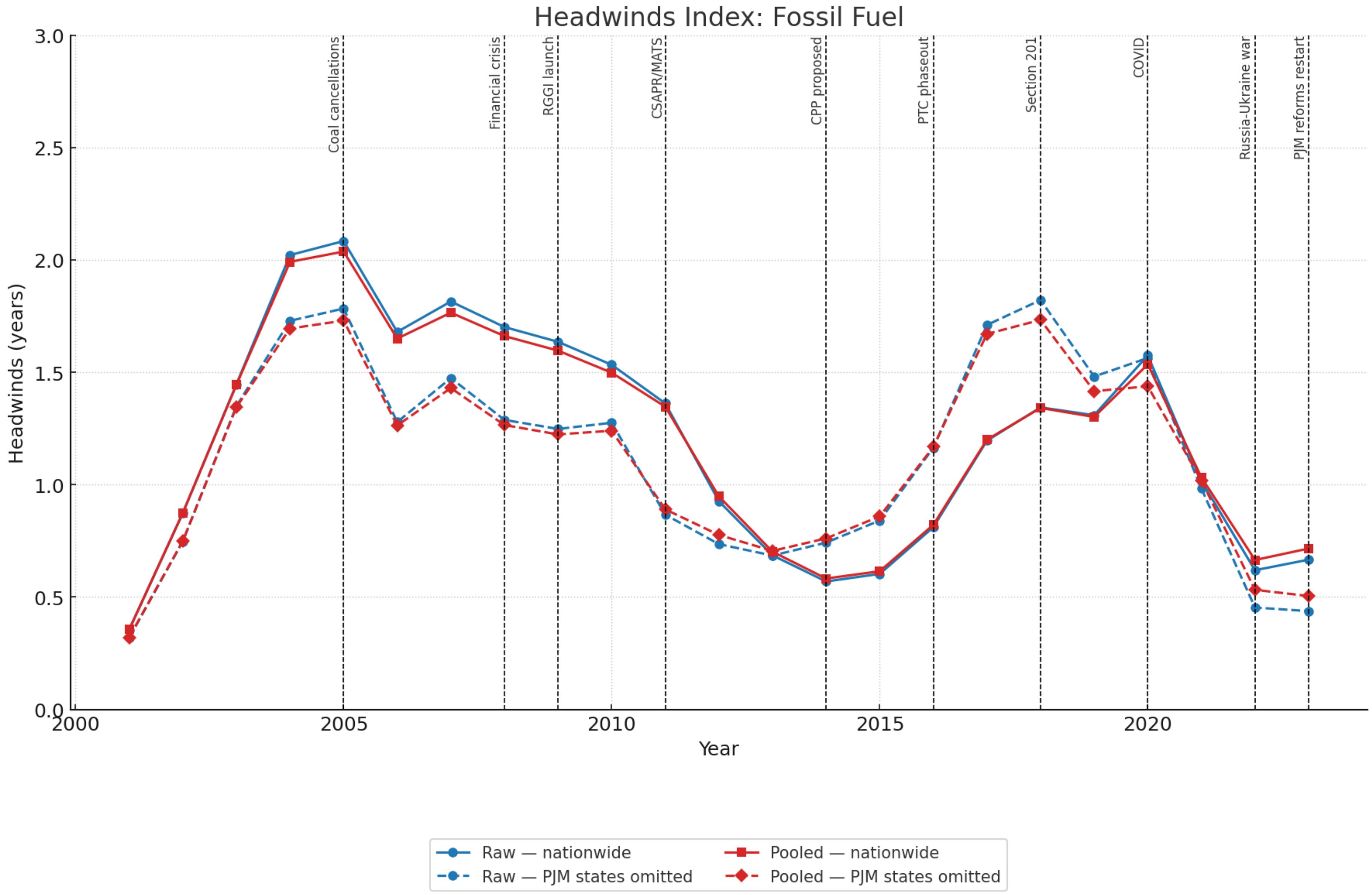

Figure 9: The plot shows Headwinds for fossil fuels (coal, natural gas, and oil).

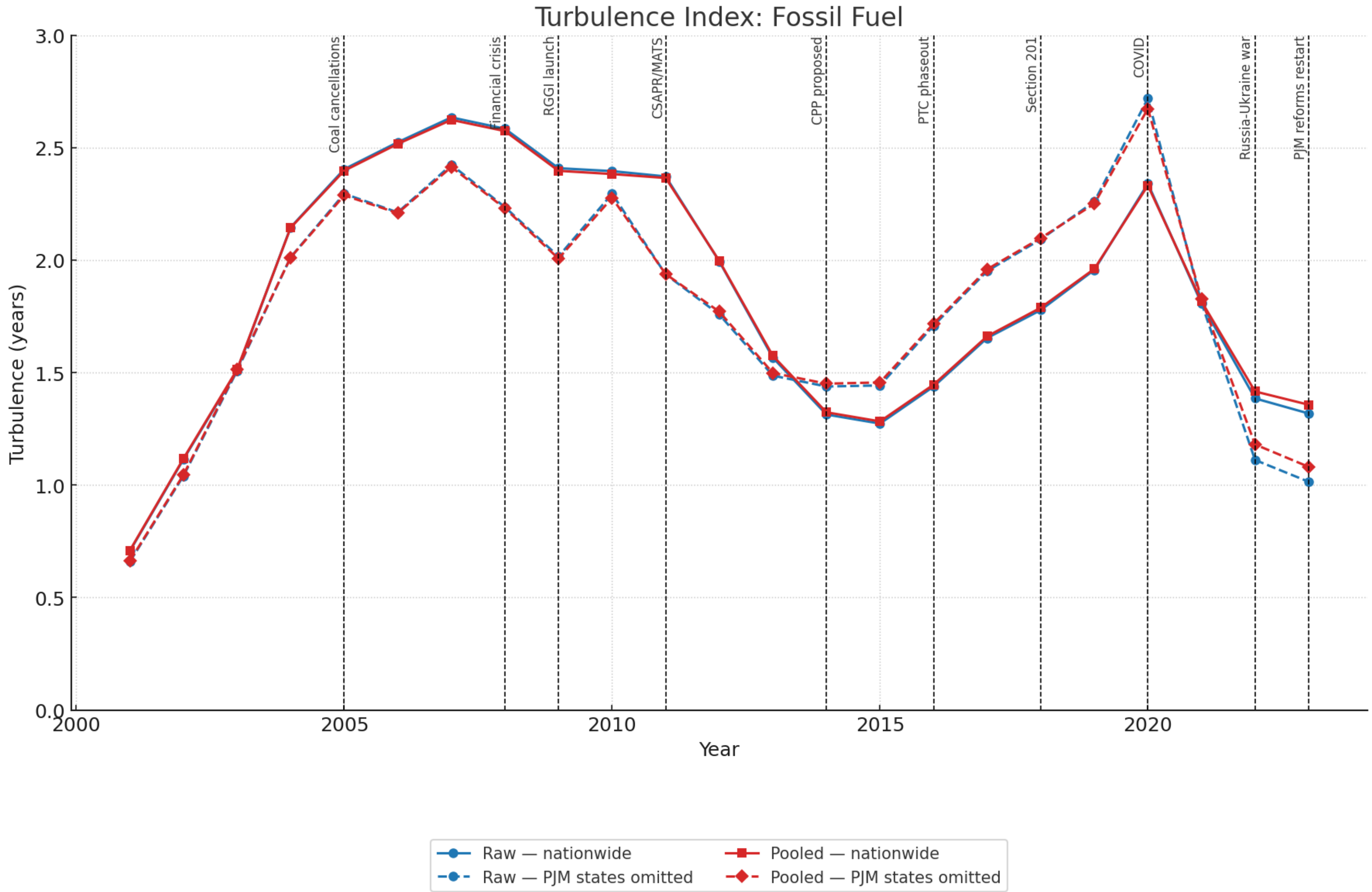

Figure 10: The plot shows Turbulence for fossil fuels (coal, natural gas, and oil).

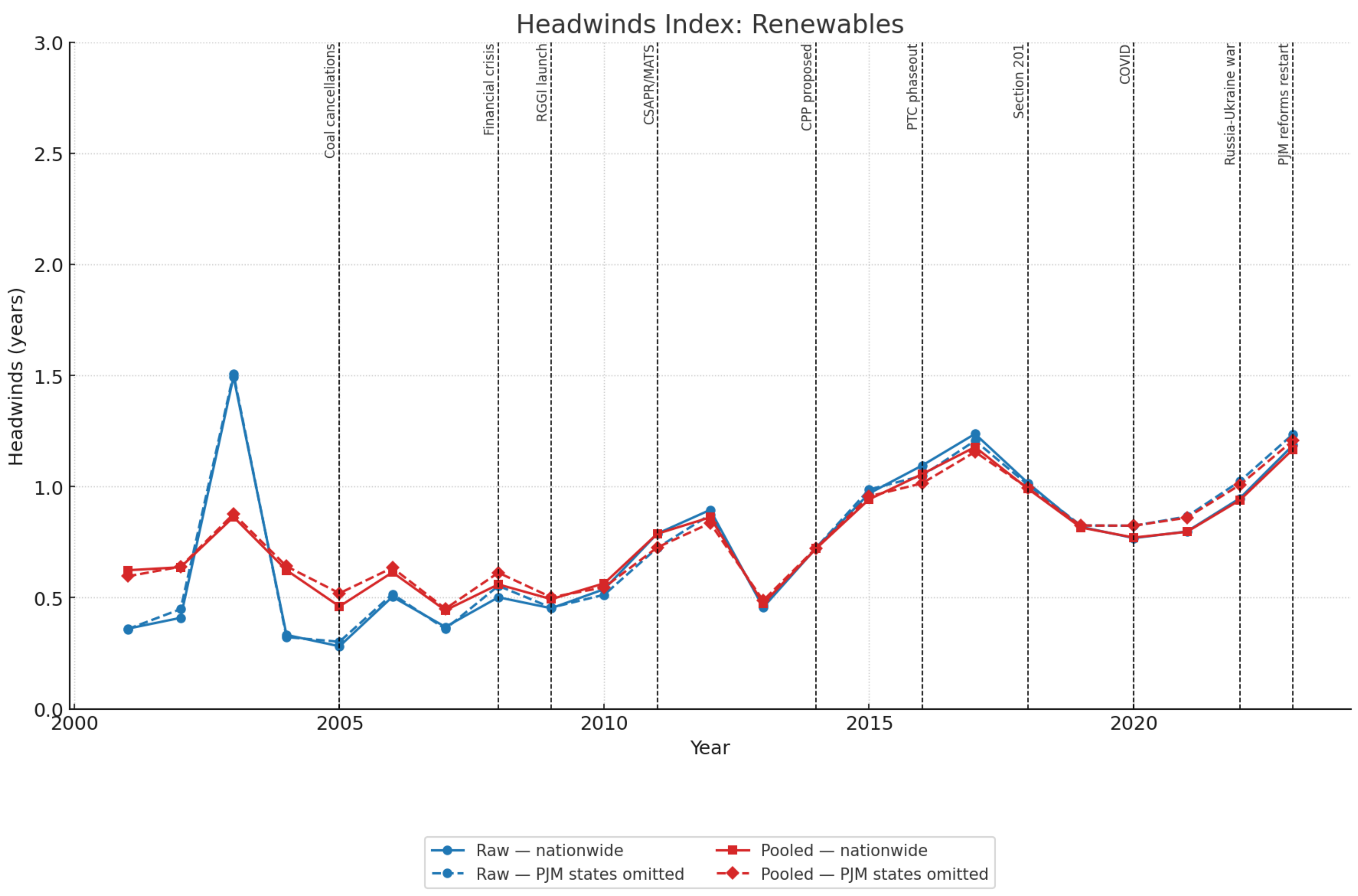

Figure 11: The plot shows Headwinds for renewable fuels (sun, wind, and water).

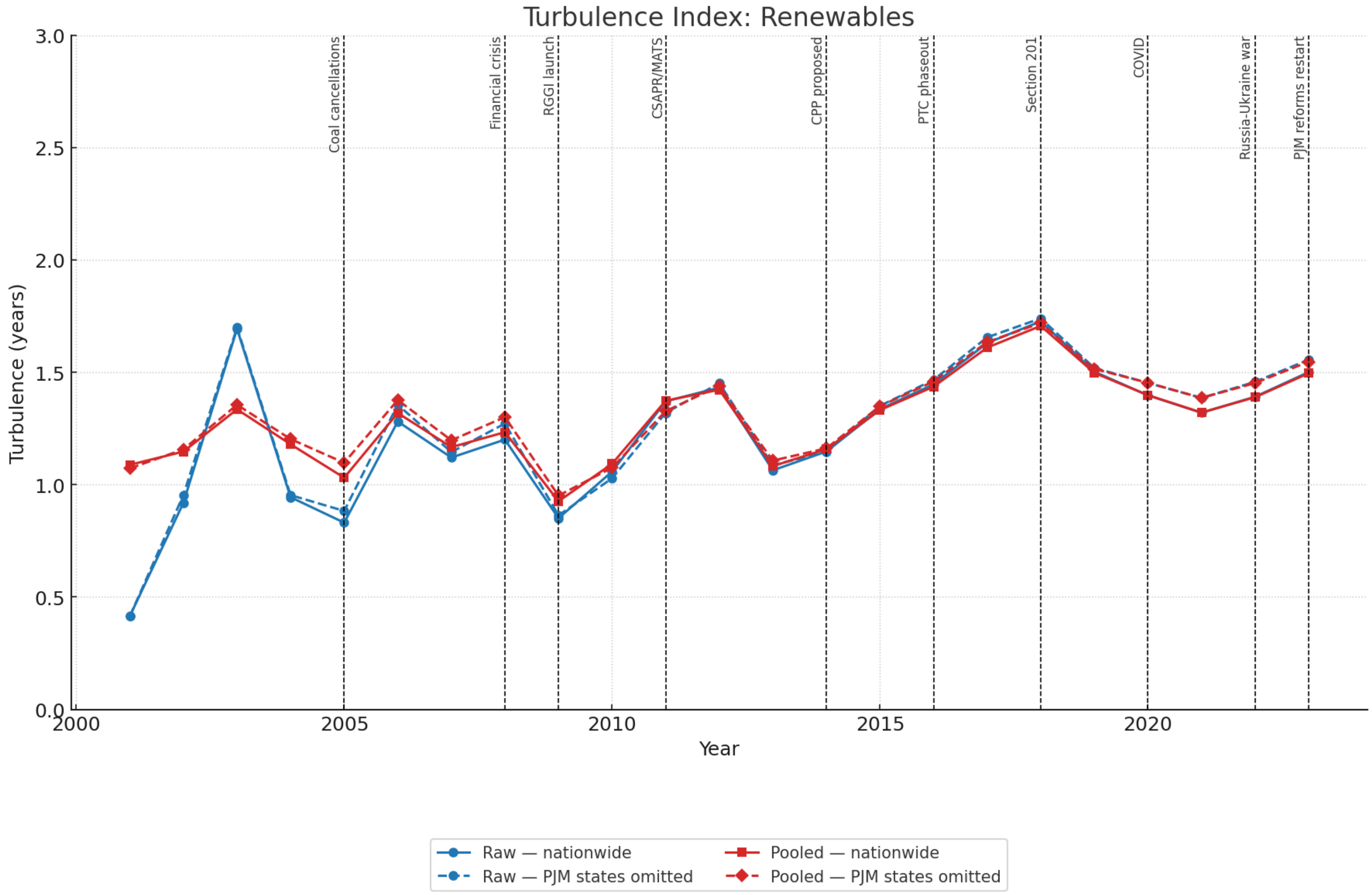

Figure 12: The plot shows Turbulence for renewable fuels (sun, wind, and water).