

# Package ‘fsPDA’

April 29, 2021

**Type** Package

**Title** Forward-Selected Panel Data Approach for Program Evaluation

**Version** 1.0.0

**Maintainer** Zhentao Shi <shizhentao@gmail.com>

**Description** This package automates the forward-selected PDA in Shi and Huang (2021).

**License** MIT

**Encoding** UTF-8

**LazyData** true

**Depends** R (>= 3.5.0)

**RoxygenNote** 7.1.1

**Imports** ggplot2, MASS

**Suggests** knitr, rmarkdown

**VignetteBuilder** knitr

## R topics documented:

china_import	1
est.fsPDA	2
fsPDA	3
HCW	4
plot.fsPDA	4
<b>Index</b>	<b>6</b>

---

china_import	<i>Chinese Import Data</i>
--------------	----------------------------

---

## Description

This dataset is described in Section 5 of Shi and Huang (2021).

## Usage

```
data(china_import)
```

## Format

- **Frequency:** Monthly
- **Date Range:** 2010:Feb – 2015:Dec
- treated Time series of luxury watch import, under the official category name "watches with case of, or clad with, precious metal". Names are in the format "yyyymm".
- control A matrix of commodities as the control units. The rowname is "yyyymm", and the column name is the identity of the commodity.
- intervention\_time Character. The month when the treatment intervention started, formatted in "yyyymm".

## Source

United Nations Comtrade Database (<https://comtrade.un.org/>). This database provides detailed statistics for international commodity trade.

Pre-treatment: 2010:Feb – 2012:Dec (35 months). Post-treatment: 2013:Jan – 2015:Dec (36 observations).

China's monthly import data are unavailable after September of 2012. To make our data consistent, we sum the other countries' export values to China, and take them as import values before and after September of 2012.

## References

Zhentao Shi and Jingyi Huang (2021): "Forward-Selected Panel Data Approach for Program Evaluation," forthcoming at the Journal of Econometrics, arXiv: 1908.05894

---

est.fsPDA

*Forward-Selected Panel Data Approach*

---

## Description

Estimation and inference by the forward-selected panel data approach (Shi and Huang, 2021).

## Usage

```
est.fsPDA(treated, control, treatment_start, date = NULL, lrvlag = NULL)
```

## Arguments

treated	Numeric. T-dimensional vector of time series of the treated unit.
control	Numeric. T-by-N matrix. Each column is the times series of a control unit.
treatment_start	An integer specifying the period treatment starts.
date	Date or numeric. A T-by-1 vector of date class or any meaningful numerical sequence. The default option NULL means 1:length(treated) is used.
lrvlag	A non-negative integer specifying the maximum lag with the Bartlett kernel for the Newey-West long-run variance estimator. The default option NULL specifies $\text{floor}((\text{length}(\text{treated}) - \text{treatment\_start} + 1)^{1/4})$ . The admissible maximum is $\text{floor}(\sqrt{\text{length}(\text{treated}) - \text{treatment\_start} + 1})$ .

**Value**

A list of the class "fsPDA" is returned. It contains the following components:

<code>select</code>	A list containing the selected control units, where <code>dim</code> is the number of selected units, <code>control</code> stores the identities of the selected units, <code>coef</code> contains the coefficient estimates, and <code>RSquared</code> is the in-sample R-Squared.
<code>in_sample</code>	A data frame of the in-sample fitted values.
<code>out_of_sample</code>	A data frame of the out-of-sample counterfactual predictions, and the estimated treatment effect (= realization - counterfactual).
<code>ATE</code>	A numeric vector containing estimate of average treatment effect (ATE), its long-run standard error, t-statistic, and p-value if the hypothesis testing of zero ATE.

**References**

Zhentao Shi and Jingyi Huang (2021): "Forward-Selected Panel Data Approach for Program Evaluation," forthcoming at the Journal of Econometrics, arXiv: 1908.05894

**Examples**

```
library(fsPDA)

# Example of China luxury watch import
data("china_import")
date_import <- names(china_import$treated)
result <- est.fsPDA(treated = china_import$treated,
                   control = china_import$control,
                   treatment_start = which(date_import == china_import$intervention_time),
                   date = as.Date(paste(substr(date_import,1,4), "-",
                                             substr(date_import, 5, 6), "-01", sep="")))

print(result)
plot(result, tlab = "Year", ylab = "Monthly Growth Rate")

# Example of HCW
data("HCW")
result <- est.fsPDA(treated = HCW$panel[,1],
                   control = HCW$panel[,-1],
                   treatment_start = HCW$T1+1,
                   date = as.Date(paste(substr(HCW$quarter, 1, 4), "-",
                                             (as.numeric(substr(HCW$quarter, 6, 6))-1)*3+1, "-1", sep="")))

print(result)
plot(result, tlab = "Year", ylab = "Real GDP Growth Rate")
```

**Description**

Automated estimation and inference by the forward-selected panel data approach (Shi and Huang, 2021).

## Details

est.fsPDA is the main function. It generates a fsPDA object which can work with the generic method plot.

## References

Zhentao Shi and Jingyi Huang (2021): "Forward-Selected Panel Data Approach for Program Evaluation," forthcoming at the Journal of Econometrics, arXiv: 1908.05894

---

HCW	<i>Hsiao, Ching and Wan (2012) Dataset</i>
-----	--------------------------------------------

---

## Description

This is the real GDP growth rate data used in Hsiao, Ching, and Wan (2012).

## Usage

```
data(HCW)
```

## Format

- panel A matrix of real GDP growth rate of 25 countries and territories, indicated by the column names. The first column is Hong Kong.
- quarter Character. A time sequence related to panel labeled by year and quarter.
- T1 An Integer indicates the length of pre-treated period.

## Source

The data archive of the original paper (<http://qed.econ.queensu.ca/jae/2012-v27.5/hsiao-ching-wan/>).

## References

Hsiao, C., S. H. Ching, and S. K. Wan (2012): "A panel data approach for program evaluation: measuring the benefits of political and economic integration of Hong Kong with mainland China," Journal of Applied Econometrics, 27(5), 705–740.

---

plot.fsPDA	<i>Plot Raw Data, Fitting, and Counterfactual Prediction</i>
------------	--------------------------------------------------------------

---

## Description

The generic plot method for objects of the "fsPDA" class It draws the observed time series of the treated unit in a red solid line. It uses the fsPDA method to fit the pre-treatment sample (represented by a green dash line), and then to predict the counterfactual in the post-treatment sample (represented by a blue dash line).

**Usage**

```
## S3 method for class 'fsPDA'
plot(
  x,
  tlab = NULL,
  ylab = NULL,
  title = NULL,
  point = TRUE,
  legend.position = "bottom"
)
```

**Arguments**

x	An object of the "fsPDA" class.
tlab	The label of the time (horizontal) axis.
ylab	The label of the value (vertical) axis.
title	The text for the title.
point	Logical. Should a layer of points be included in the plot?
legend.position	The position of legends ("none", "left", "right", "bottom", "top").

**Note**

"ggplot2" package must be installed as dependency.

**See Also**

[est.fsPDA](#) for examples.

# Index

## \* datasets

- china\_import, [1](#)
- HCW, [4](#)

china\_import, [1](#)

est.fsPDA, [2](#), [5](#)

fsPDA, [3](#)

HCW, [4](#)

plot.fsPDA, [4](#)