CSSS 594 / POLS 559:

Time Series and Panel Data for the Social Sciences

Panel Data Models with Many Time Periods

Christopher Adolph

Department of Political Science
and
Center for Statistics and the Social Sciences
University of Washington, Seattle

Estimating Panel Models

Last time, we discussed how including random and/or fixed effects changes the properties of our estimators of β

In this lecture, we'll talk about how to estimate and interpet panel models using fixed and/or random effects

And how to decide if we need (or even can use) fixed effects

We can always add random effects, but in some cases FEs either be too costly to estimate (in terms of dfs), or simply impossible to estimate

Estimating Panel Models

Last time, we discussed how including random and/or fixed effects changes the properties of our estimators of β

In this lecture, we'll talk about how to estimate and interpet panel models using fixed and/or random effects

And how to decide if we need (or even can use) fixed effects

We can always add random effects, but in some cases FEs either be too costly to estimate (in terms of dfs), or simply impossible to estimate

We will consider first the small N, large T case,
which allows more complex time series modeling

Then the large N, small T case
which raises the possibility of bias in fixed effects estimation

Finally, we consider heteroskedasticity in time or across panel structures

Estimating Fixed Effects Models

Option 1: Fixed effects or "within" estimator:

$$y_{it} - \bar{y}_i = (\mathrm{x}_{it} - \bar{\mathrm{x}}_i)\boldsymbol{\beta} + (u_{it} - \bar{u}_{it})$$

- estimating the fixed effects by differencing them out

- including time-invariant variables directly in x_{it} impossible here

- (there are workarounds, e.g., if we have an instrument for the time-invariant variable that is uncorrelated with the fixed effects; see Hausman-Taylor)

- suggests a complementary "between" estimator of $\bar{y}_i$ on $\bar{x}_i$ which could include time-invariant x_i; together these models explain the variance in y_{it}

- does not actually provide estimates of the fixed effects themselves; just purges them from the model to remove omitted time-invariant variables

Estimating Fixed Effects Models

Option 2: Dummy variable estimator (sometimes called LSDV)

$$y_{it} = \mathrm{x}_{it}\boldsymbol{\beta} + \alpha_i + u_{it}$$

- yields estimates of α_i fixed effects (may be useful in quest for omitted variables; see if the α_i look like a variable you know)

- for large T, should be very similar to FE estimator

- not a good idea for very small T: estimates of α_i will be poor

Time-Invariant Covariates & Fixed Effects

We can't include time-invariant variables in fixed effects models

If we do, we will have perfect collinearity, and can't get estimates

That is, we will get some parameter estimates equal to `NA`

Never report a regression with `NA` parameters

The regression you tried to run was impossible. Start over with a possible one.

Time-Invariant Covariates & Fixed Effects

If we can't include time-invariant variables in a fixed effects model,
does that mean time-invariant variables can never explain changes over time?

Time-Invariant Covariates & Fixed Effects

If we can't include time-invariant variables in a fixed effects model,
does that mean time-invariant variables can never explain changes over time?

You might think so: how can a constant explain a variable?

But time-invariant variables could still effect time-varying outcomes in a special way. . .

Time-Invariant Covariates & Fixed Effects

If we can't include time-invariant variables in a fixed effects model, does that mean time-invariant variables can never explain changes over time?

You might think so: how can a constant explain a variable?

But time-invariant variables could still effect time-varying outcomes in a special way. . .

Time-invariant variables can influence how a unit weathers time-varying *shocks* in some other variable

Example: labor market regulations (e.g. employment protection) don't change much over time

Blanchard & Wolfers found that when a negative economic shock hits, unemployment may rebound more slowly where such protections are stronger

Time-Invariant Covariates & Fixed Effects

We can model how a slow moving or time-invariant covariate conditions the effect of a quickly changing covariate on y_{it}

To estimate how a time-invariant covariate x_{it} mediates the effect of a shock, s_{it}, include on the RHS $x_{it} \times s_{it}$ and s_{it}, while omitting x_{it} itself

(It's okay *and necessary* to omit the x_{it} base term in this special case, because α_i already captures the effect of x_{it})

Many theories about institutions can be tested this way

Time-Invariant Covariates & Fixed Effects

What if we want to "include" time-invariant covariates' effect on the long term average level of y?

We might partition the fixed effect into:

1. the portion "explained" by known time-invariant variables and
2. the portion still unexplained

Plümper & Troeger have methods to do this.

In this case, our estimates of the time-invariant effects are vulnerable to omitted variable bias from unmeasured time-invariant variables, even though time varying variables in the model are not

Thus you now need to control for *lots* of time-invariant variables directly, even hard to measure ones like culture

Estimating Random & Mixed Effects Models

Estimation of random effects is by maximum likelihood (ML)
or generalized least squares (GLS)

Technically we're just adding one parameter to estimate:
the variance of the random effects, σ_α^2

This is partitioned out of the overall variance, σ^2

Can understand this most easily by abstracting away from time series for a moment

Estimating Random & Mixed Effects Models

Recall that for linear regression, we assume homoskedastic, serially uncorrelated errors, and thus a variance-covariance matrix like this:

$$\Omega = \begin{bmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & \sigma^2 \end{bmatrix}$$

Estimating Random & Mixed Effects Models

And recall that heteroskedastic (but serially uncorrelated) errors have this variance-covariance matrix

$$\Omega = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{bmatrix}$$

Estimating Random & Mixed Effects Models

And finally, remember heteroskedastic, serially correlated errors follow this general form of variance-covariance

$$\Omega = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2 \end{bmatrix}$$

What does this matrix look like for random effects with no serial correlation?

Estimating Random & Mixed Effects Models

Define the variance of the random effect as

$$\mathrm{E}(\alpha_i^2) = \sigma_\alpha^2 = \mathrm{var}(\alpha_i)$$

Estimating Random & Mixed Effects Models

Define the variance of the random effect as

$$\mathrm{E}(\alpha_i^2) = \sigma_\alpha^2 = \mathrm{var}(\alpha_i)$$

Define the expected value of the squared white noise term as σ_ε^2

$$\mathrm{E}(\varepsilon_{it}^2) = \sigma_\varepsilon^2 = \mathrm{var}(\varepsilon_{it})$$

Estimating Random & Mixed Effects Models

Define the variance of the random effect as

$$\mathrm{E}(\alpha_i^2) = \sigma_\alpha^2 = \mathrm{var}(\alpha_i)$$

Define the expected value of the squared white noise term as σ_ε^2

$$\mathrm{E}(\varepsilon_{it}^2) = \sigma_\varepsilon^2 = \mathrm{var}(\varepsilon_{it})$$

White noise is serially uncorrelated, so has covariance 0 for $t \neq s$:

$$\mathrm{E}(\varepsilon_{it}\varepsilon_{is}) = 0 = \mathrm{cov}(\varepsilon_{it}, \varepsilon_{is})$$

Estimating Random & Mixed Effects Models

Define the variance of the random effect as

$$\mathrm{E}(\alpha_i^2) = \sigma_\alpha^2 = \mathrm{var}(\alpha_i)$$

Define the expected value of the squared white noise term as σ_ε^2

$$\mathrm{E}(\varepsilon_{it}^2) = \sigma_\varepsilon^2 = \mathrm{var}(\varepsilon_{it})$$

White noise is serially uncorrelated, so has covariance 0 for $t \neq s$:

$$\mathrm{E}(\varepsilon_{it}\varepsilon_{is}) = 0 = \mathrm{cov}(\varepsilon_{it}, \varepsilon_{is})$$

Finally, note that we assumed the white noise error and random effect are uncorrelated,

$$\mathrm{E}(\alpha_i\varepsilon_{it}) = 0 = \mathrm{cov}(\alpha_i, \varepsilon_{it})$$

Estimating Random & Mixed Effects Models

Thus the variance of the whole random component of the model is

$$\mathrm{E}((\alpha_i + \varepsilon_{it})(\alpha_i + \varepsilon_{it})) = \mathrm{E}(\alpha_i^2) + 2\mathrm{E}(\alpha_i \varepsilon_{it}) + \mathrm{E}(\varepsilon_{it}^2)$$

Estimating Random & Mixed Effects Models

Thus the variance of the whole random component of the model is

$$\begin{aligned}\mathrm{E}((\alpha_i + \varepsilon_{it})(\alpha_i + \varepsilon_{it})) &= \mathrm{E}(\alpha_i^2) + 2\mathrm{E}(\alpha_i \varepsilon_{it}) + \mathrm{E}(\varepsilon_{it}^2) \\ &= \sigma_\alpha^2 + 0 + \sigma_\varepsilon^2\end{aligned}$$

Estimating Random & Mixed Effects Models

Thus the variance of the whole random component of the model is

$$\begin{aligned}
\mathrm{E}((\alpha_i + \varepsilon_{it})(\alpha_i + \varepsilon_{it})) &= \mathrm{E}(\alpha_i^2) + 2\mathrm{E}(\alpha_i \varepsilon_{it}) + \mathrm{E}(\varepsilon_{it}^2) \\
&= \sigma_\alpha^2 + 0 + \sigma_\varepsilon^2 \\
&= \sigma_\alpha^2 + \sigma_\varepsilon^2
\end{aligned}$$

Estimating Random & Mixed Effects Models

Thus the variance of the whole random component of the model is

$$
\begin{aligned}
\mathrm{E}((\alpha_i + \varepsilon_{it})(\alpha_i + \varepsilon_{it})) &= \mathrm{E}(\alpha_i^2) + 2\mathrm{E}(\alpha_i \varepsilon_{it}) + \mathrm{E}(\varepsilon_{it}^2) \\
&= \sigma_\alpha^2 + 0 + \sigma_\varepsilon^2 \\
&= \sigma_\alpha^2 + \sigma_\varepsilon^2
\end{aligned}
$$

And the covariance of the whole random component is:

$$
\mathrm{E}((\alpha_i + \varepsilon_{it})(\alpha_i + \varepsilon_{is})) = \mathrm{E}(\alpha_i^2) + \mathrm{E}(\alpha_i \varepsilon_{is}) + \mathrm{E}(\alpha_i \varepsilon_{it}) + \mathrm{E}(\varepsilon_{it} \varepsilon_{is})
$$

Estimating Random & Mixed Effects Models

Thus the variance of the whole random component of the model is

$$\begin{aligned}
\mathrm{E}((\alpha_i + \varepsilon_{it})(\alpha_i + \varepsilon_{it})) &= \mathrm{E}(\alpha_i^2) + 2\mathrm{E}(\alpha_i \varepsilon_{it}) + \mathrm{E}(\varepsilon_{it}^2) \\
&= \sigma_\alpha^2 + 0 + \sigma_\varepsilon^2 \\
&= \sigma_\alpha^2 + \sigma_\varepsilon^2
\end{aligned}$$

And the covariance of the whole random component is:

$$\begin{aligned}
\mathrm{E}((\alpha_i + \varepsilon_{it})(\alpha_i + \varepsilon_{is})) &= \mathrm{E}(\alpha_i^2) + \mathrm{E}(\alpha_i \varepsilon_{is}) + \mathrm{E}(\alpha_i \varepsilon_{it}) + \mathrm{E}(\varepsilon_{it} \varepsilon_{is}) \\
&= \sigma_\alpha^2 + 0 + 0 + 0
\end{aligned}$$

Estimating Random & Mixed Effects Models

Thus the variance of the whole random component of the model is

$$\begin{aligned}
\mathrm{E}((\alpha_i + \varepsilon_{it})(\alpha_i + \varepsilon_{it})) &= \mathrm{E}(\alpha_i^2) + 2\mathrm{E}(\alpha_i \varepsilon_{it}) + \mathrm{E}(\varepsilon_{it}^2) \\
&= \sigma_\alpha^2 + 0 + \sigma_\varepsilon^2 \\
&= \sigma_\alpha^2 + \sigma_\varepsilon^2
\end{aligned}$$

And the covariance of the whole random component is:

$$\begin{aligned}
\mathrm{E}((\alpha_i + \varepsilon_{it})(\alpha_i + \varepsilon_{is})) &= \mathrm{E}(\alpha_i^2) + \mathrm{E}(\alpha_i \varepsilon_{is}) + \mathrm{E}(\alpha_i \varepsilon_{it}) + \mathrm{E}(\varepsilon_{it} \varepsilon_{is}) \\
&= \sigma_\alpha^2 + 0 + 0 + 0 \\
&= \sigma_\alpha^2
\end{aligned}$$

Estimating Random & Mixed Effects Models

If our data have a single random effect in the mean for each unit
$\rightarrow$ serially correlated errors, but expressable using only two variances:

- the random effects variance σ^2_α
- the white noise term's variance σ^2_ε

$$\Omega = \begin{bmatrix} \sigma^2_\alpha + \sigma^2_\varepsilon & \sigma^2_\alpha & \sigma^2_\alpha & \sigma^2_\alpha \\ \sigma^2_\alpha & \sigma^2_\alpha + \sigma^2_\varepsilon & \sigma^2_\alpha & \sigma^2_\alpha \\ \sigma^2_\alpha & \sigma^2_\alpha & \sigma^2_\alpha + \sigma^2_\varepsilon & \sigma^2_\alpha \\ \sigma^2_\alpha & \sigma^2_\alpha & \sigma^2_\alpha & \sigma^2_\alpha + \sigma^2_\varepsilon \end{bmatrix}$$

Estimating Random & Mixed Effects Models

We have drastically simplified this matrix, and can now use FGLS (Feasible Generalized Least Squares) or ML to estimate it

$$\hat{\boldsymbol{\beta}}_{\text{GLS}} = \left(\sum_{i=1}^{N} \mathrm{X}_i' \Omega^{-1} \mathrm{X}_i \right)^{-1} \left(\sum_{i=1}^{N} \mathrm{X}_i' \hat{\Omega}^{-1} \mathrm{y}_i \right)$$

where X_i is the $T \times K$ matrix of covariates for unit i, all times $t = 1, \ldots T$, and all K covariates

All we need are the estimates $\hat{\sigma}_\alpha^2$ and $\hat{\sigma}_\varepsilon^2$, and we can calculate $\hat{\boldsymbol{\beta}}_{\text{GLS}}$

Estimating Random & Mixed Effects Models

We get $\hat{\sigma}_\varepsilon^2$ from the residuals from a LS regression:

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{NT - K} \sum_{i=1}^{N} \sum_{t=1}^{T} \hat{\varepsilon}_{it\mathrm{LS}}^2$$

(This is the usual estimator, but for NT observations)

Estimating Random & Mixed Effects Models

To get an estimator of $\hat{\sigma}^2_\alpha$, we need to adjust for the fact that we have only so many unique pairs of errors to compare:

$$\hat{\sigma}^2_\alpha = \mathrm{E}\left(\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}(\alpha_i + \varepsilon_{it})(\alpha_i + \varepsilon_{is})\right)$$

Estimating Random & Mixed Effects Models

To get an estimator of $\hat{\sigma}^2_\alpha$, we need to adjust for the fact that we have only so many unique pairs of errors to compare:

$$\hat{\sigma}^2_\alpha = \mathrm{E}\left(\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}(\alpha_i + \varepsilon_{it})(\alpha_i + \varepsilon_{is})\right)$$

$$= \mathrm{E}\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}((\alpha_i + \varepsilon_{it})(\alpha_i + \varepsilon_{is}))$$

Estimating Random & Mixed Effects Models

To get an estimator of $\hat{\sigma}^2_\alpha$, we need to adjust for the fact that we have only so many unique pairs of errors to compare:

$$\begin{aligned}
\hat{\sigma}^2_\alpha &= \mathrm{E}\left(\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}(\alpha_i + \varepsilon_{it})(\alpha_i + \varepsilon_{is})\right) \\
&= \mathrm{E}\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}((\alpha_i + \varepsilon_{it})(\alpha_i + \varepsilon_{is})) \\
&= \sum_{t=1}^{T-1}\sum_{s=t+1}^{T}\sigma^2_\alpha
\end{aligned}$$

Estimating Random & Mixed Effects Models

To get an estimator of $\hat{\sigma}^2_\alpha$, we need to adjust for the fact that we have only so many unique pairs of errors to compare:

$$\begin{aligned}
\hat{\sigma}^2_\alpha &= \mathrm{E}\left(\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}(\alpha_i+\varepsilon_{it})(\alpha_i+\varepsilon_{is})\right)\\
&= \mathrm{E}\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}\left((\alpha_i+\varepsilon_{it})(\alpha_i+\varepsilon_{is})\right)\\
&= \sum_{t=1}^{T-1}\sum_{s=t+1}^{T}\sigma^2_\alpha\\
&= \sigma^2_\alpha\sum_{t=1}^{T-1}(T-t)
\end{aligned}$$

Estimating Random & Mixed Effects Models

To get an estimator of $\hat{\sigma}_\alpha^2$, we need to adjust for the fact that we have only so many unique pairs of errors to compare:

$$\begin{aligned}
\hat{\sigma}_\alpha^2 &= \mathrm{E}\left(\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}(\alpha_i+\varepsilon_{it})(\alpha_i+\varepsilon_{is})\right) \\
&= \mathrm{E}\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}((\alpha_i+\varepsilon_{it})(\alpha_i+\varepsilon_{is})) \\
&= \sum_{t=1}^{T-1}\sum_{s=t+1}^{T}\sigma_\alpha^2 \\
&= \sigma_\alpha^2\sum_{t=1}^{T-1}(T-t) \\
&= \sigma_\alpha^2((T-1)+(T-2)+\ldots+2+1)
\end{aligned}$$

Estimating Random & Mixed Effects Models

$$= \sigma_\alpha^2((T-1) + (T-2) + \ldots + 2 + 1)$$

Estimating Random & Mixed Effects Models

$$
\begin{aligned}
&= \sigma_\alpha^2((T-1)+(T-2)+\ldots+2+1) \\
&= \sigma_\alpha^2 T(T-1)/2
\end{aligned}
$$

Estimating Random & Mixed Effects Models

$$= \sigma_\alpha^2((T-1)+(T-2)+\ldots+2+1)$$
$$= \sigma_\alpha^2 T(T-1)/2$$
$$\hat{\sigma}_\alpha^2 = \frac{1}{NT(T-1)/2-K}\sum_{i=1}^{N}\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}\hat{\varepsilon}_{it}\hat{\varepsilon}_{is}$$

where in the last step we replace σ_α^2 with its estimator from pooled LS (the average of the products of the unique pairs of residuals)

Estimating Random & Mixed Effects Models

$$
\begin{aligned}
&= \sigma_\alpha^2((T-1)+(T-2)+\ldots+2+1) \\
&= \sigma_\alpha^2 T(T-1)/2
\end{aligned}
$$

$$
\hat{\sigma}_\alpha^2 = \frac{1}{NT(T-1)/2 - K} \sum_{i=1}^{N} \sum_{t=1}^{T-1} \sum_{s=t+1}^{T} \hat{\varepsilon}_{it} \hat{\varepsilon}_{is}
$$

where in the last step we replace σ_α^2 with its estimator from pooled LS (the average of the products of the unique pairs of residuals)

With some algebra, this approach extends to serial correlaton of other kinds (ARMA)

For complex models, with many levels and/or hyperparameters, best to go Bayesian, set diffuse priors on the parameters, and use MCMC

Selecting Fixed Effects vs Random Effects Models

Choosing random effects when α_i is actually correlated with x_{it}
will lead to omitted variable bias

Choosing fixed effects when α_i is really uncorrelated with x_{it}
will lead to inefficient estimates of β (compared to random effects estimation)
and kick out our time-invariant variables

Often in comparative we are certain there are important omitted time invariant variables (culture, unmeasured institutions, long effects of history)

So choice to include fixed effects requires nothing more than theory

Still could include random effects in addition to the fixed effects

Selecting Fixed Effects vs Random Effects Models

But if we are uncertain, or want to check against estimating unnecessary fixed effects, we can use the Hausman test for (any) fixed effexts versus just having random effecxts

Hausman sets up the null hypothesis of random effects

Attempts to reject it in favor of fixed effects

Selecting Fixed Effects vs Random Effects Models

But if we are uncertain, or want to check against estimating unnecessary fixed effects, we can use the Hausman test for (any) fixed effexts versus just having random effecxts

Hausman sets up the null hypothesis of random effects

Attempts to reject it in favor of fixed effects

Checks whether the random α_i's are correlated with $\mathrm{x_i}$ under the null

Does this by calculating the variance-covariance matrices of regressors under FE and then just RE

Null is no correlation between these covariances

Selecting Fixed Effects vs Random Effects Models

But if we are uncertain, or want to check against estimating unnecessary fixed effects, we can use the Hausman test for (any) fixed effexts versus just having random effecxts

Hausman sets up the null hypothesis of random effects

Attempts to reject it in favor of fixed effects

Checks whether the random α_i's are correlated with $\mathrm{x_i}$ under the null

Does this by calculating the variance-covariance matrices of regressors under FE and then just RE

Null is no correlation between these covariances

If there is no correlation, that means the regressors do not predict the random effects (ie, are uncorrelated)

Rejecting the null suggests you may need fixed effects to deal with omitted variable bias

`phtest` in `plm` library

Interpreting Random Effects Models

Usually, interest focuses on the percentage of variance explained by the random effects

And how this variance compares to that remaining in the model

Reported by your estimation routine

What if T is very small?

If T is very small (< 15 perhaps), estimating panel dynamics efficiently and without bias gets harder

In these cases, we should investigate alternatives:

1. First differencing the series to produce a stationary, hopefully white noise processs
2. Including fixed effects for the time period (time dummies)
3. Checking for serial correlation after estimation (LM test)
4. Using lags of the dependent variable, while removing the bias from including lags with fixed effects by instrumenting with lagged differences (Arellano-Bond)

Example: GDP in a panel

Let's use the Przeworski et al democracy data to try out our variable intercept models

This exercise is for pedagogical purposes only; the models we fit are badly specified

We will investigate the following model:

$$\Delta^d \mathrm{GDP}_{it} = \alpha_i + \beta_1 \mathrm{OIL}_{it} + \beta_2 \mathrm{REG}_{it} + \beta_3 \mathrm{EDT}_{it} + \nu_{it}$$

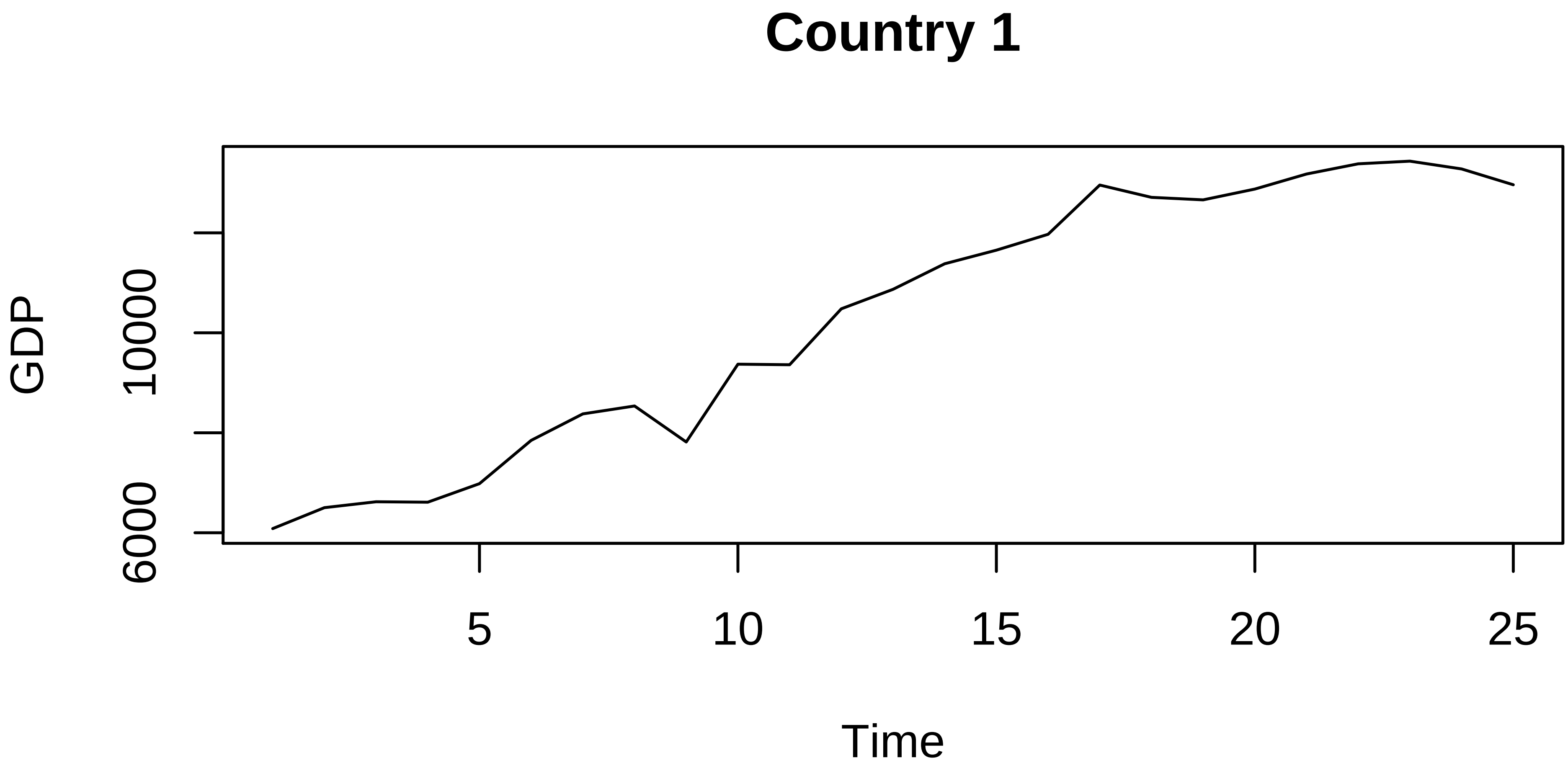
- where $\nu_{it} \sim \mathrm{ARIMA}(p, d, q)$,
- d may be 0 or 1, and
- α_i may be fixed, random, or a mixed

Example: GDP in a panel

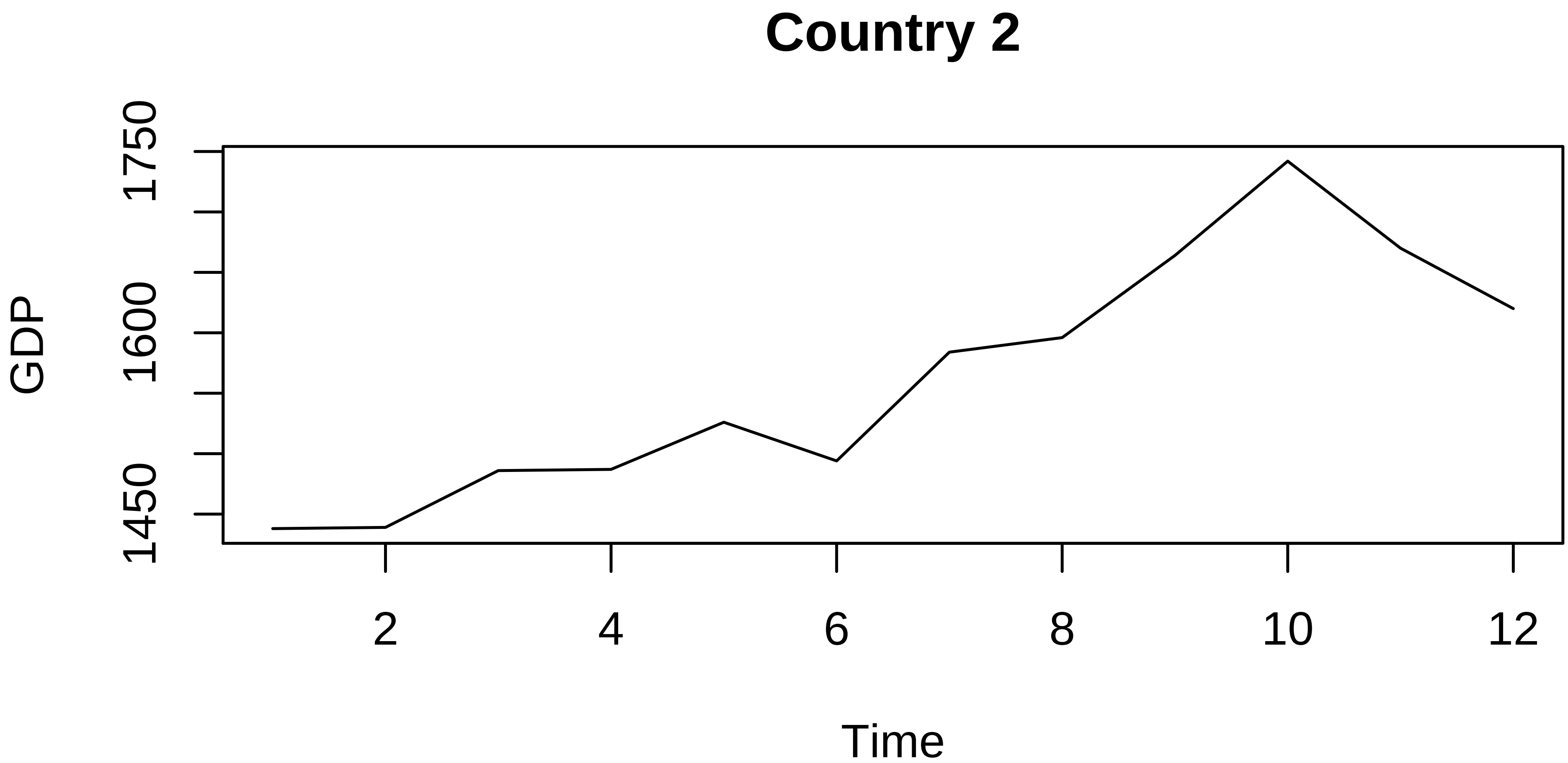
We first investigate the time series properties of GDP

But we have $N = 113$ countries! So we would have to look at 113 time series plots, 113 ACF plots, and 113 PACF plots

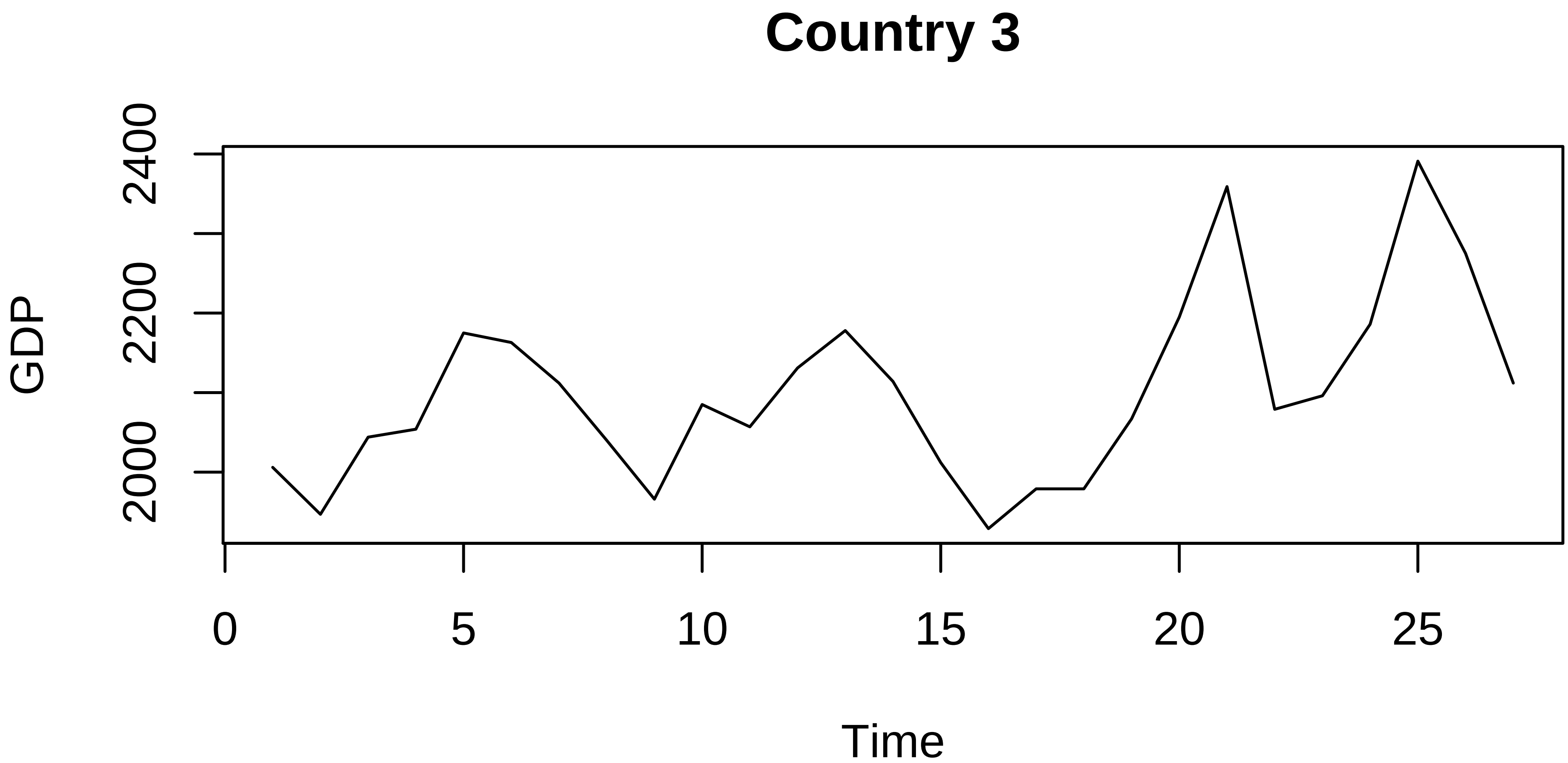
Fortunately, they do look fairly similar. . .



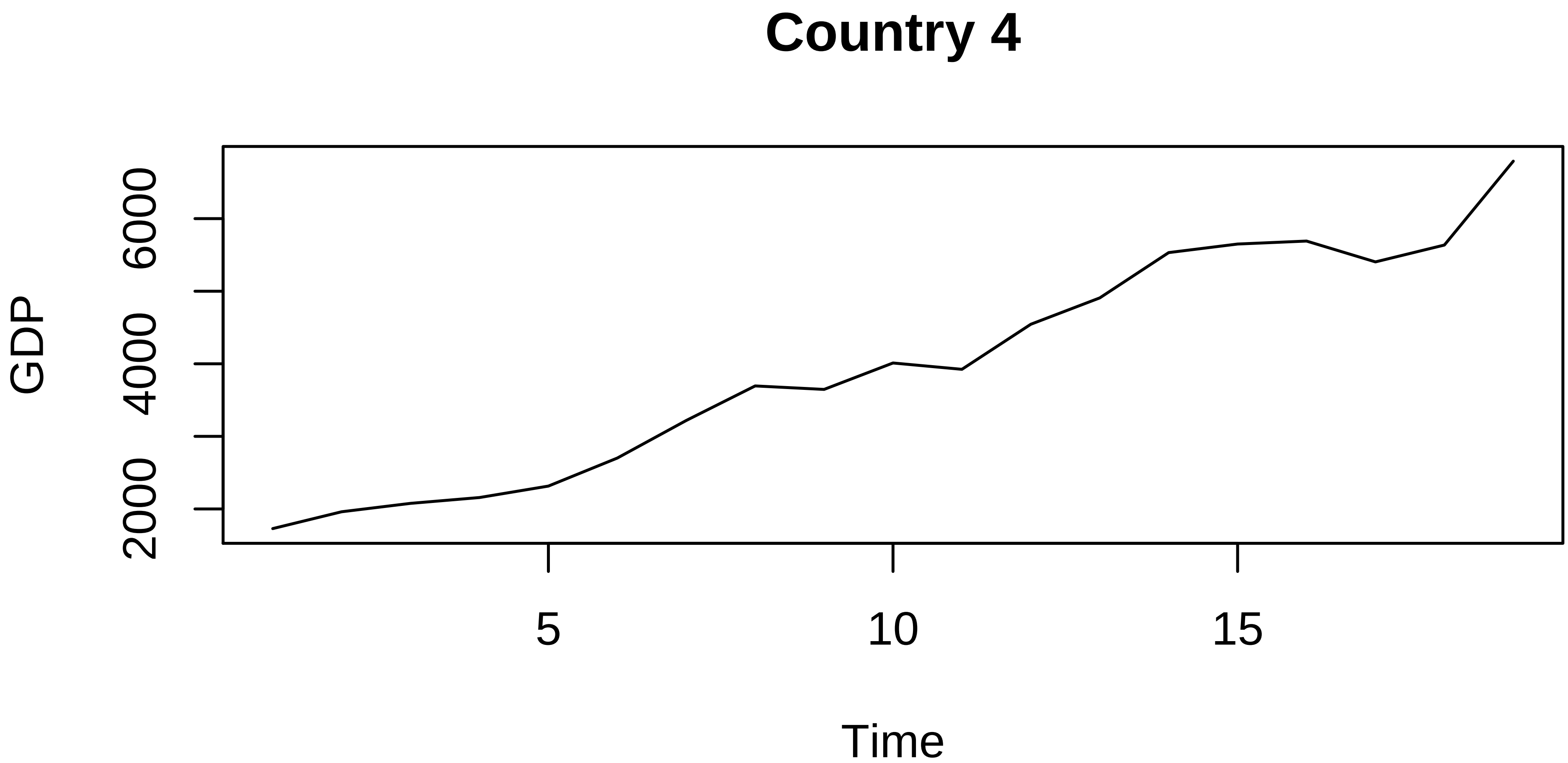
GDP time series for country 1



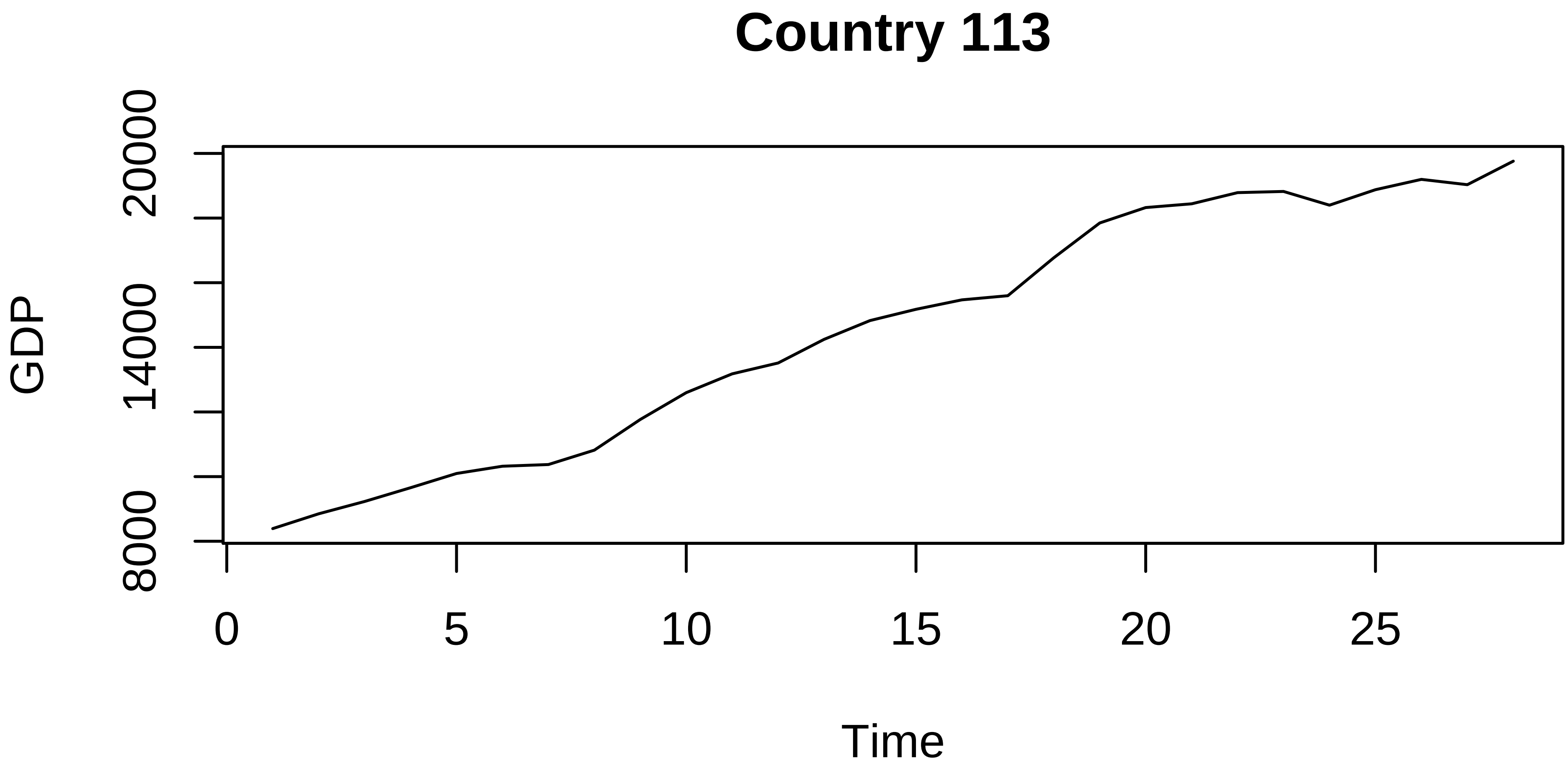
GDP time series for country 2



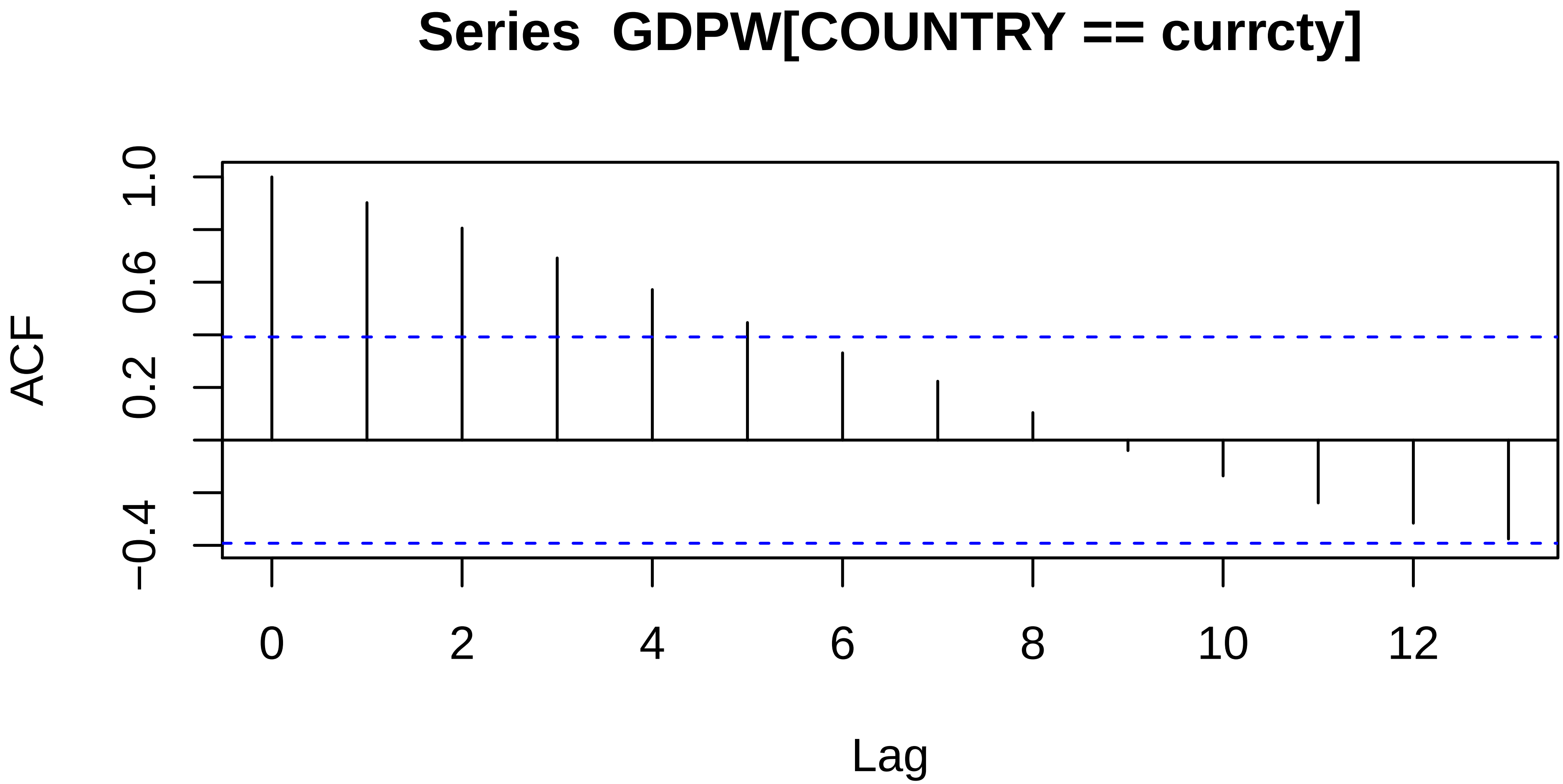
GDP time series for country 3



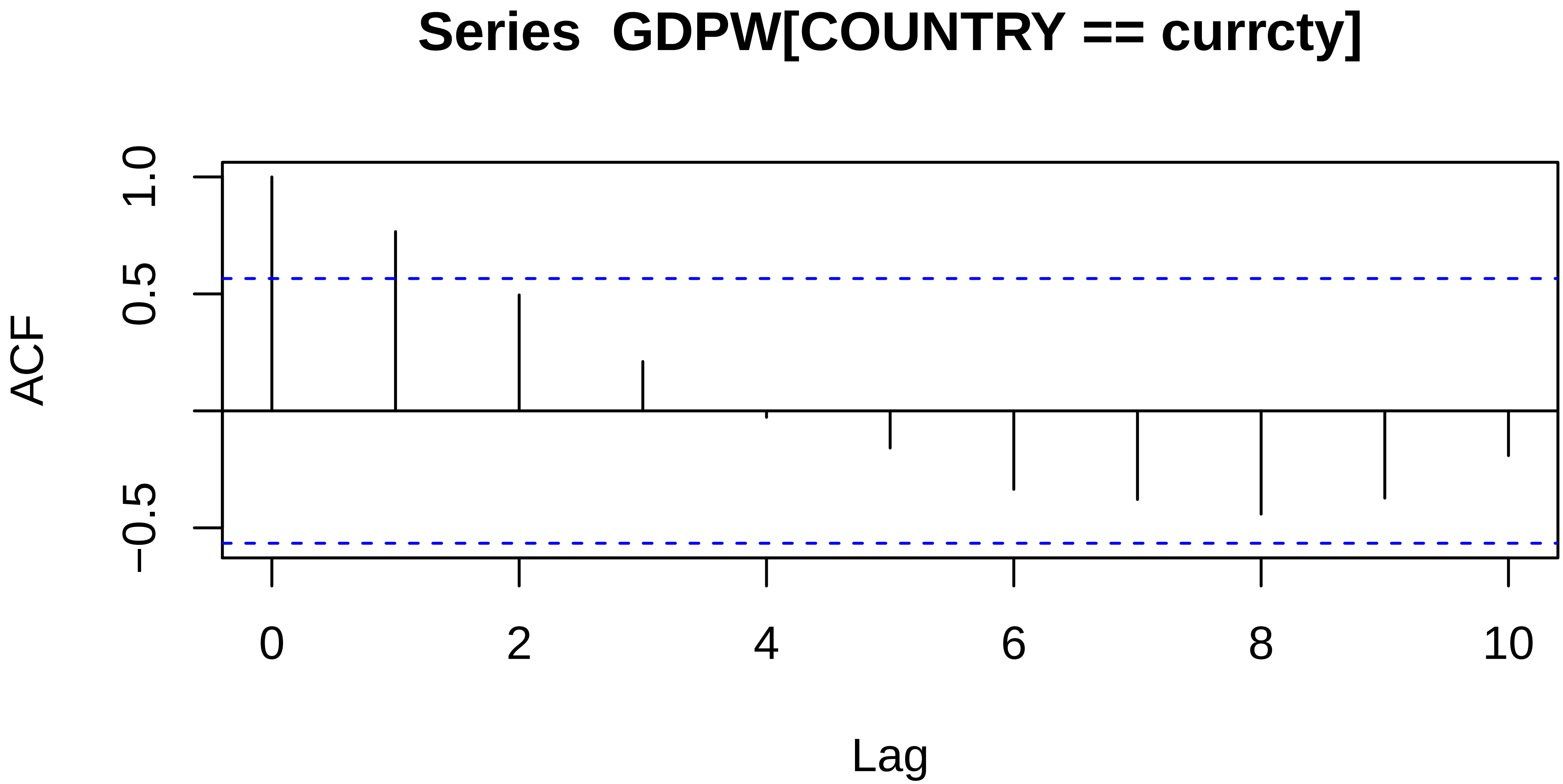
GDP time series for country 4



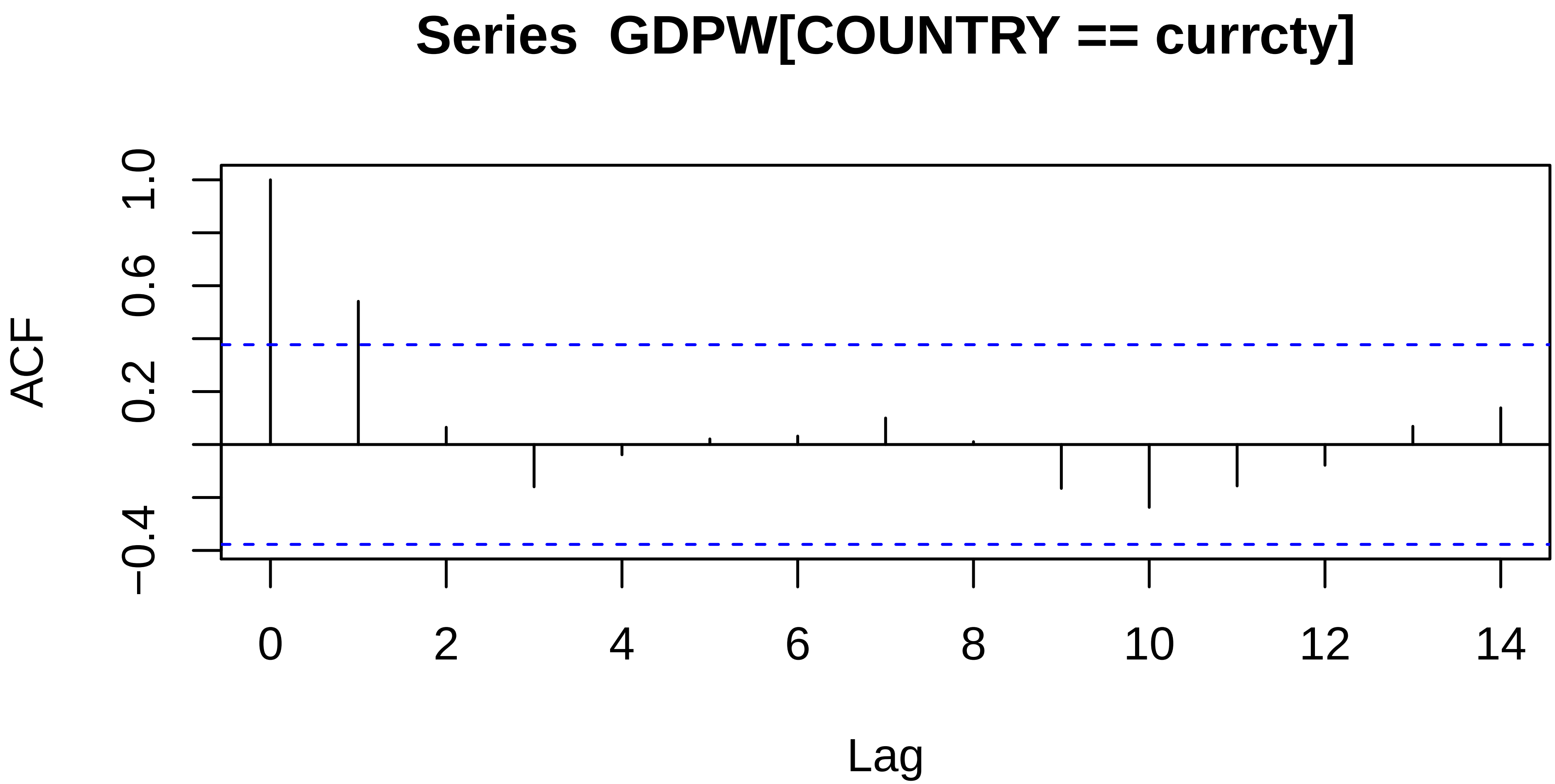
GDP time series for country 113



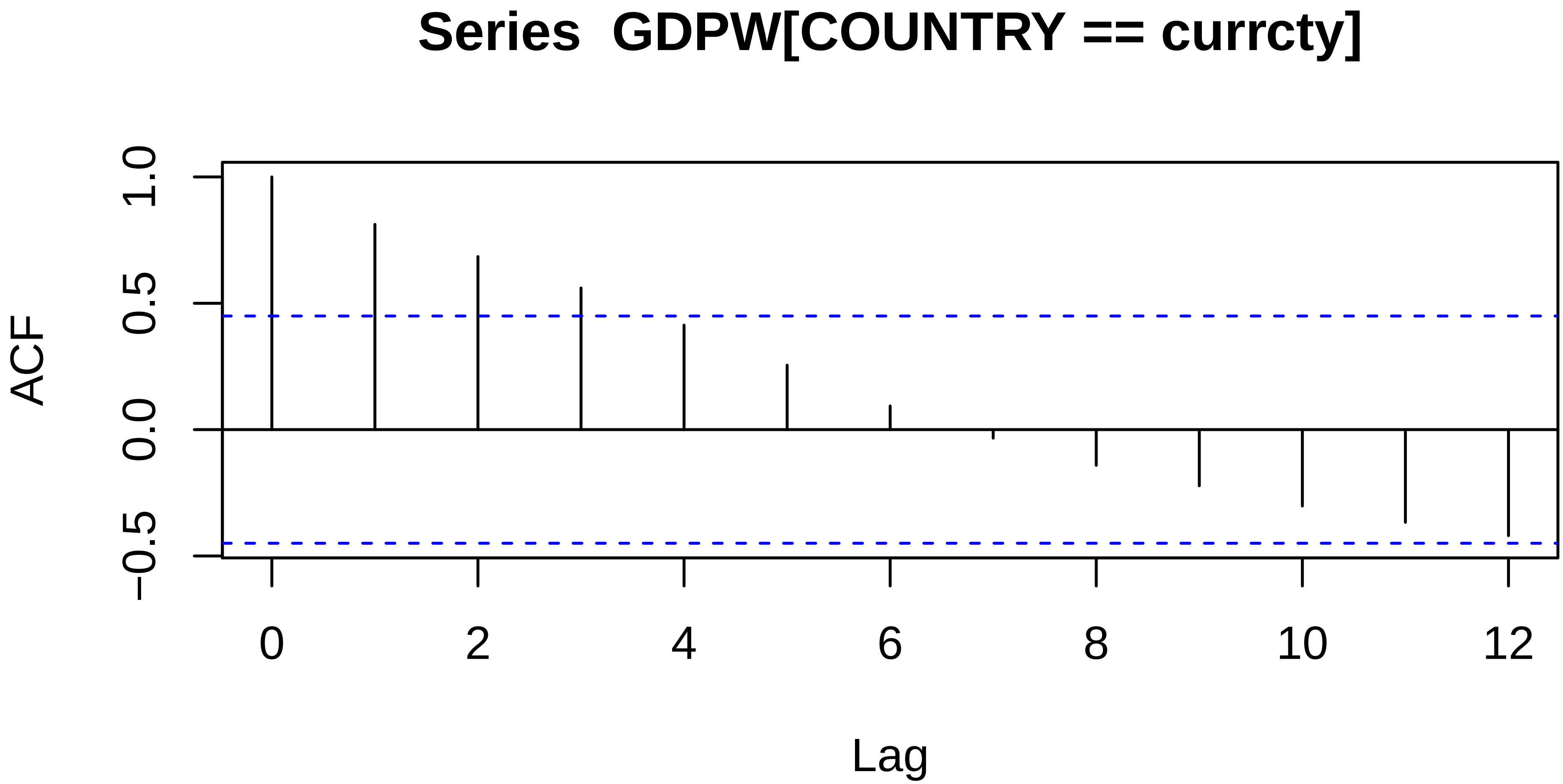
GDP ACF for country 1



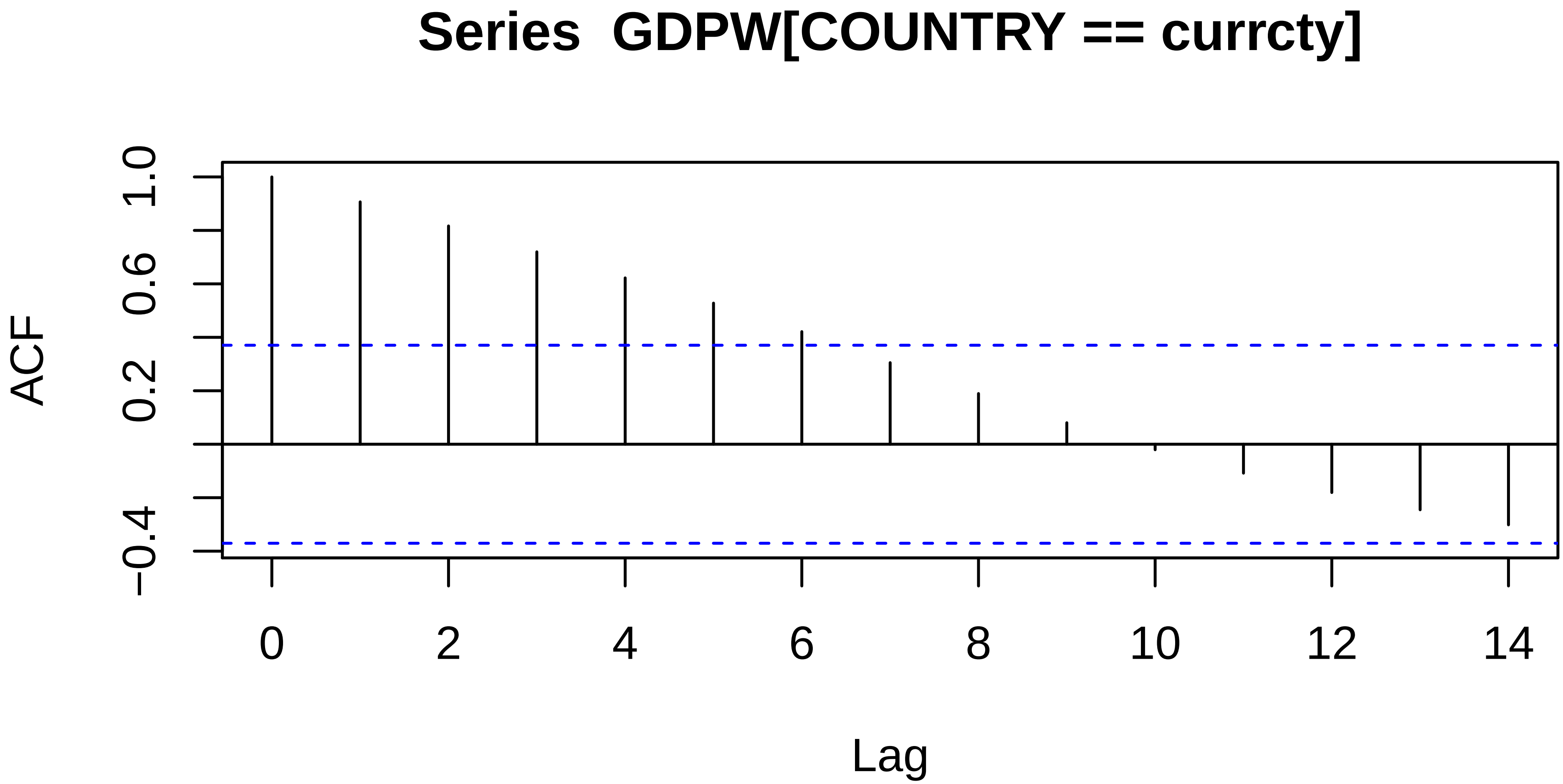
GDP ACF for country 2



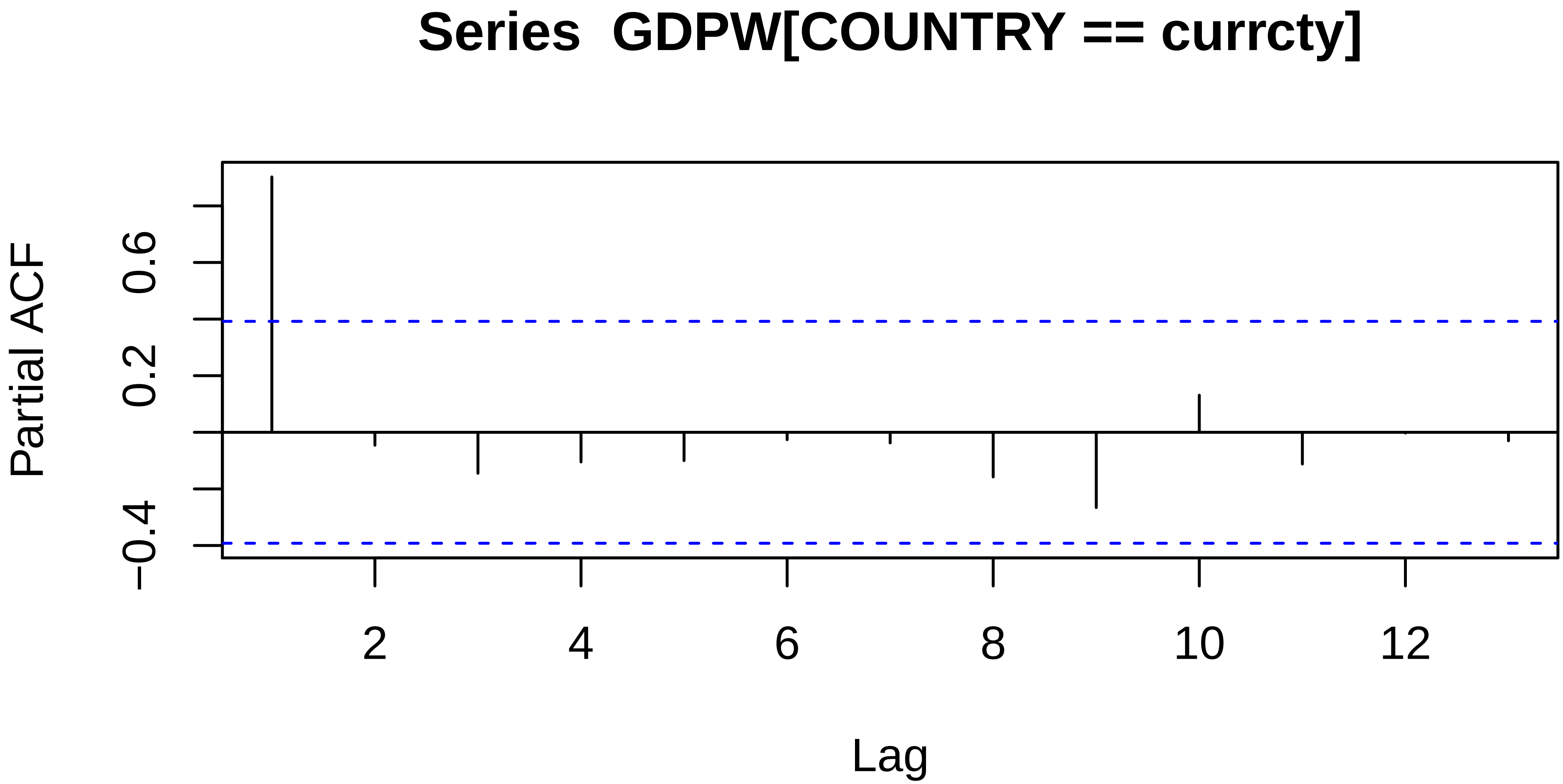
GDP ACF for country 3



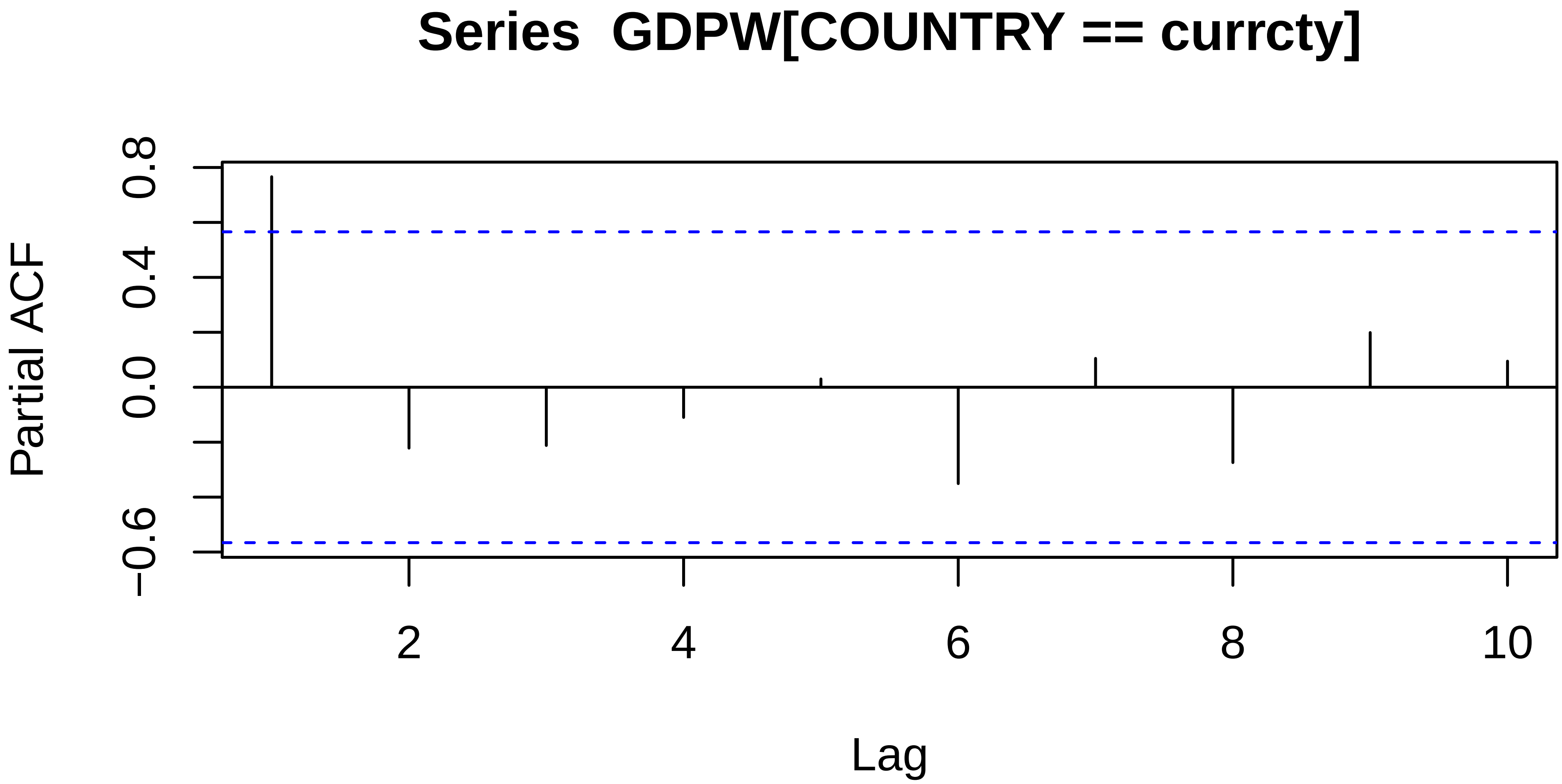
GDP ACF for country 4



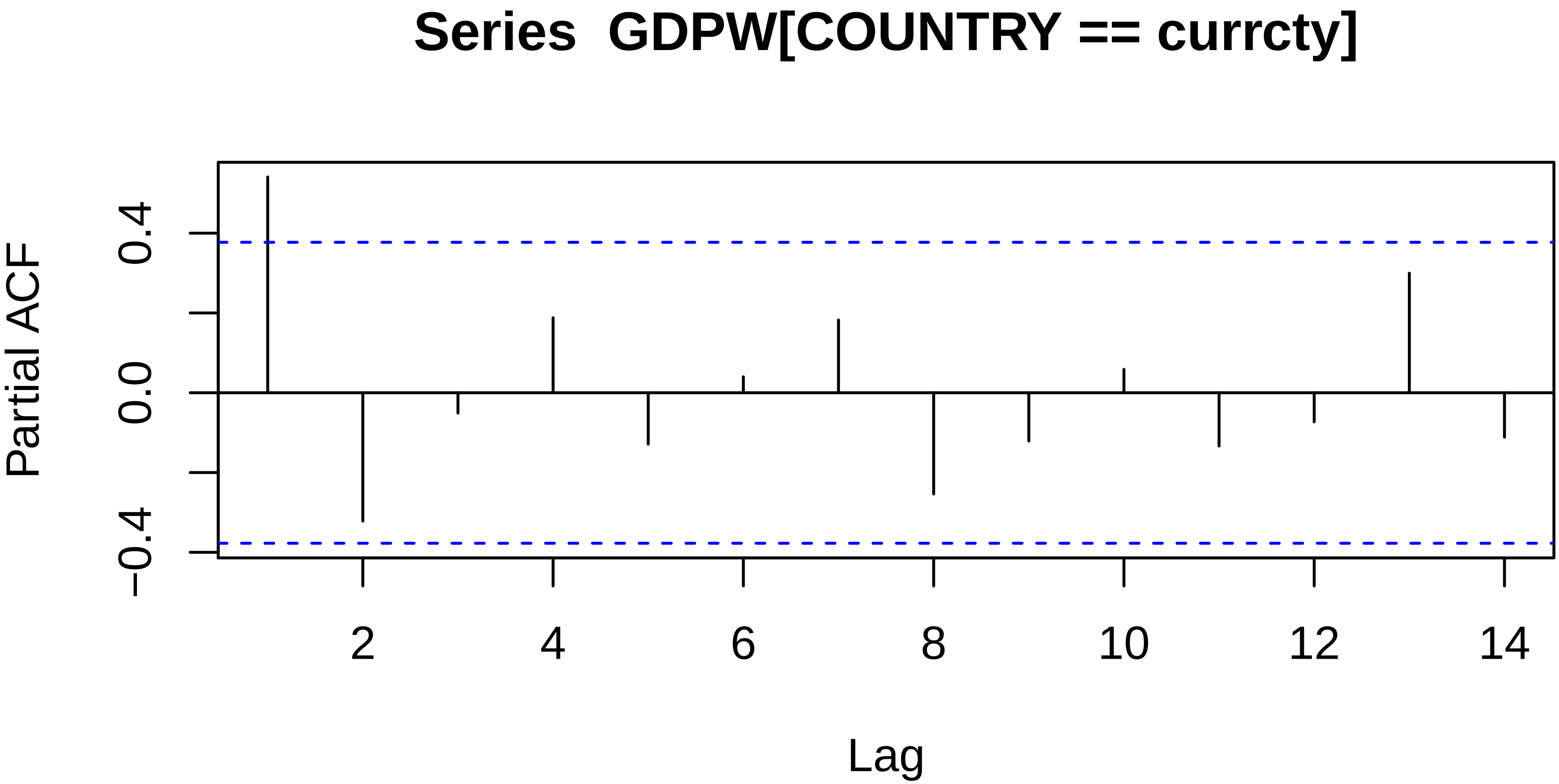
GDP ACF for country 113



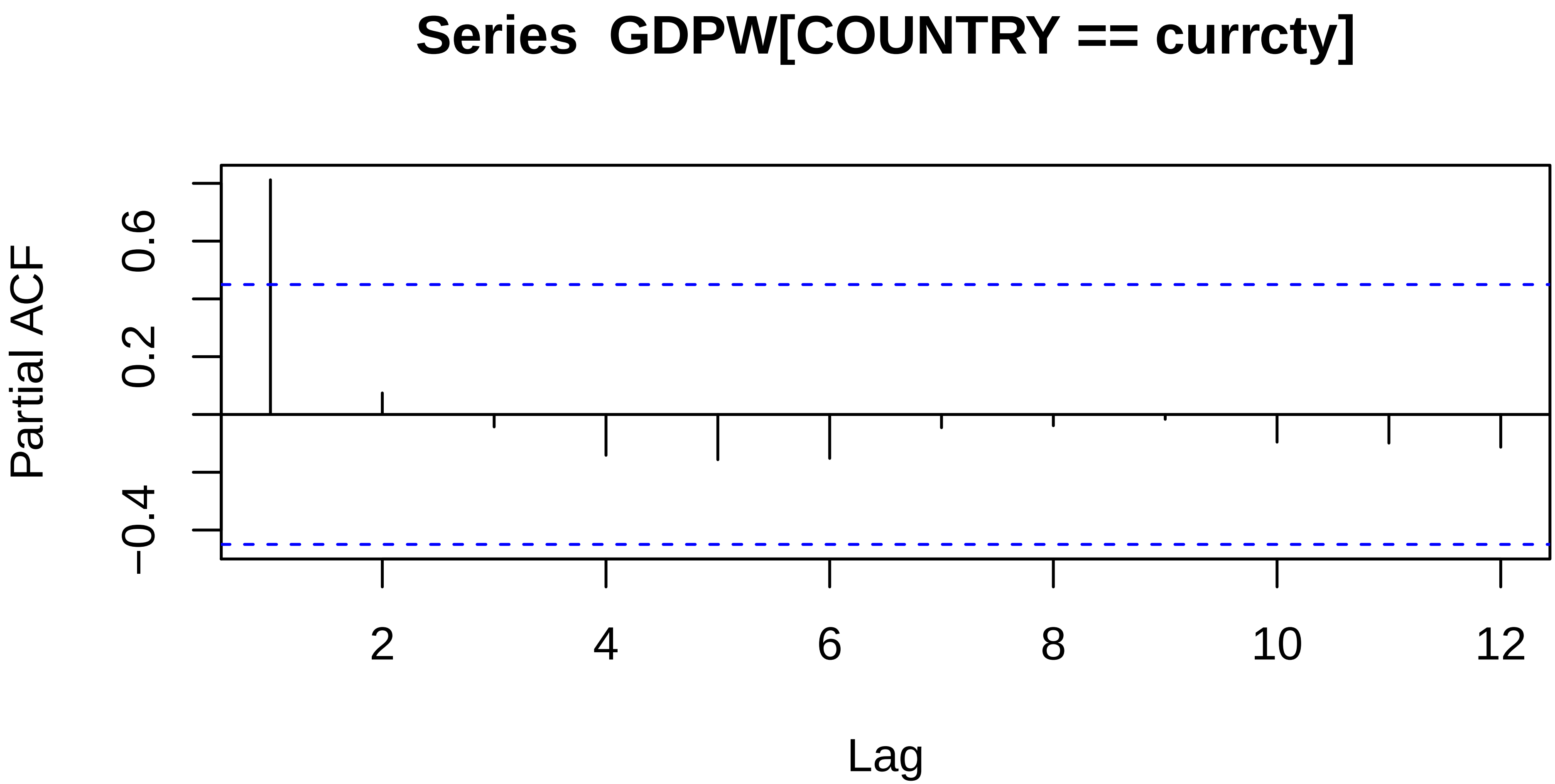
GDP PACF for country 1



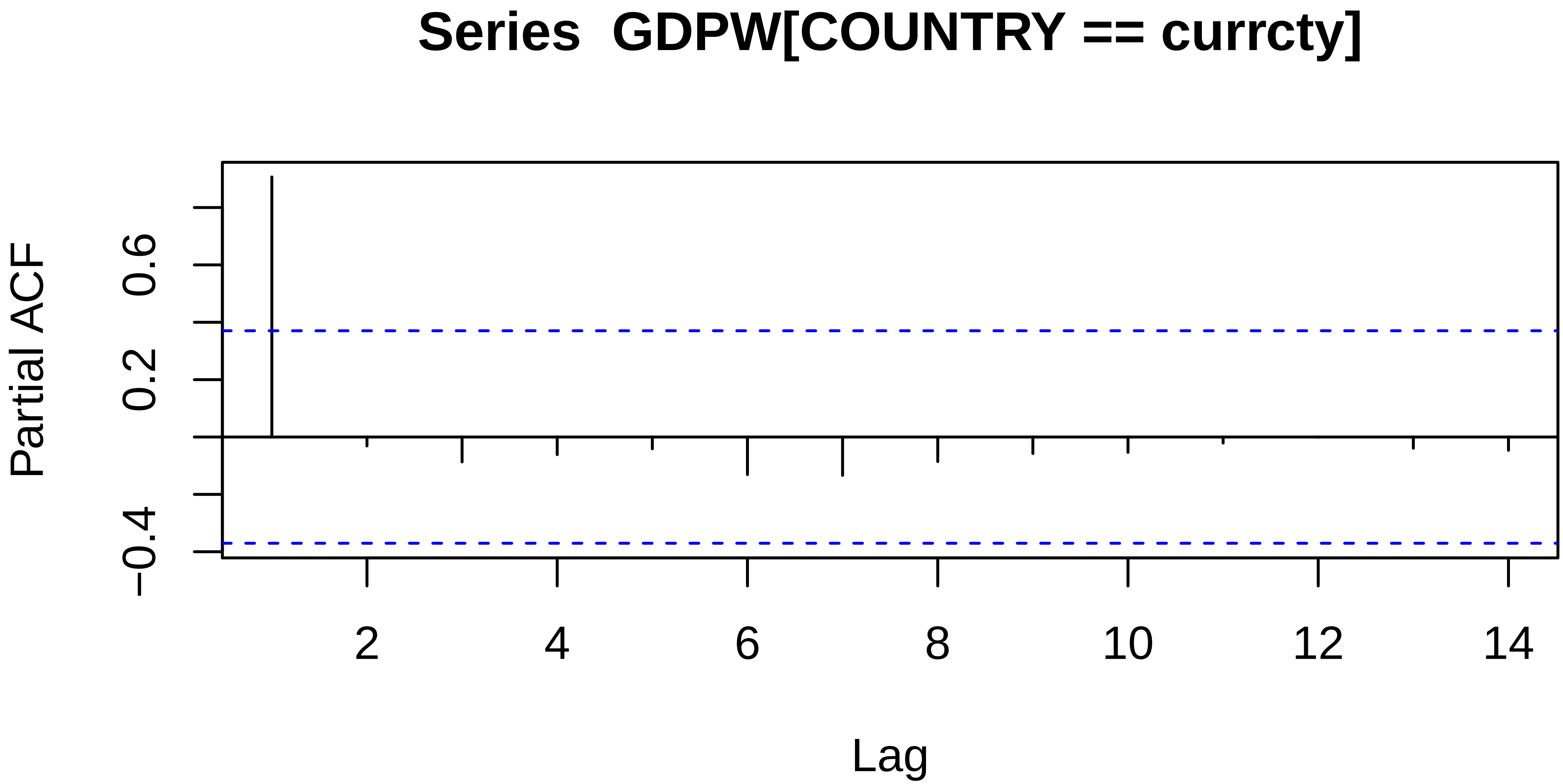
GDP PACF for country 2



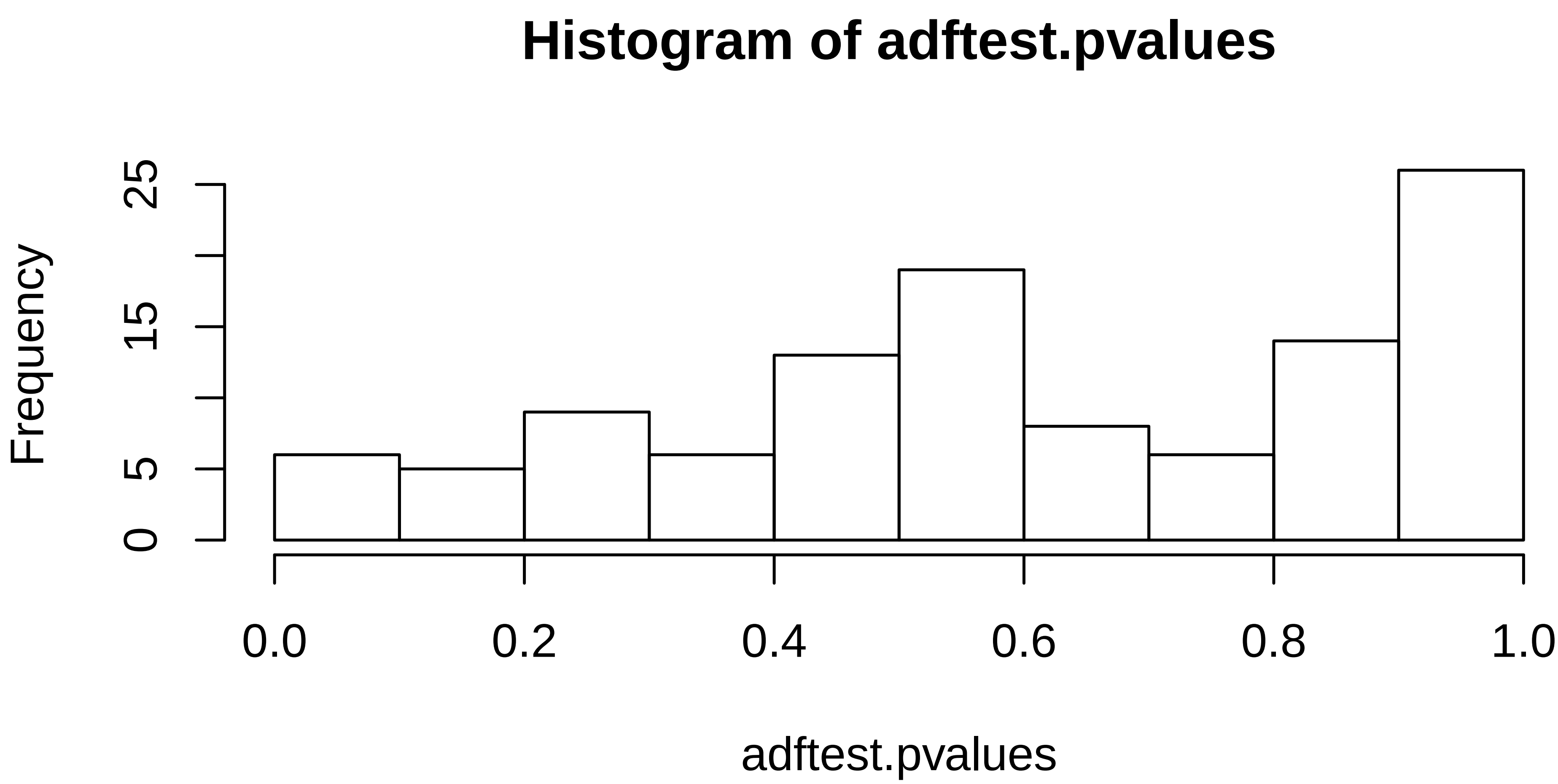
GDP PACF for country 3



GDP PACF for country 4

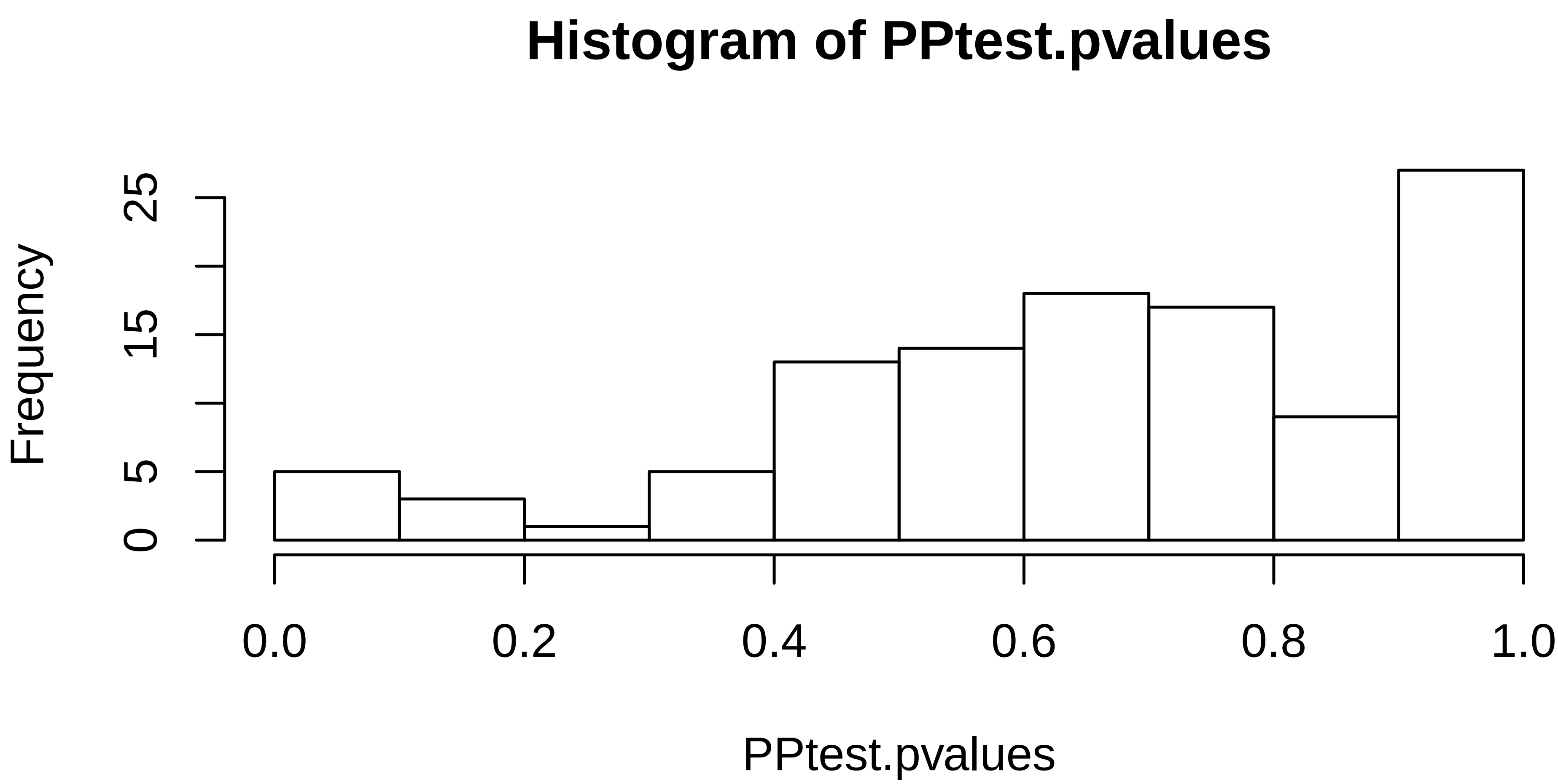


GDP PACF for country 113



Histogram of p-values from ADF tests on GDPW

What would we see if there were no unit roots?



Histogram of p-values from Phillips-Peron tests on GDPW

Choosing AR(p,q) for panel

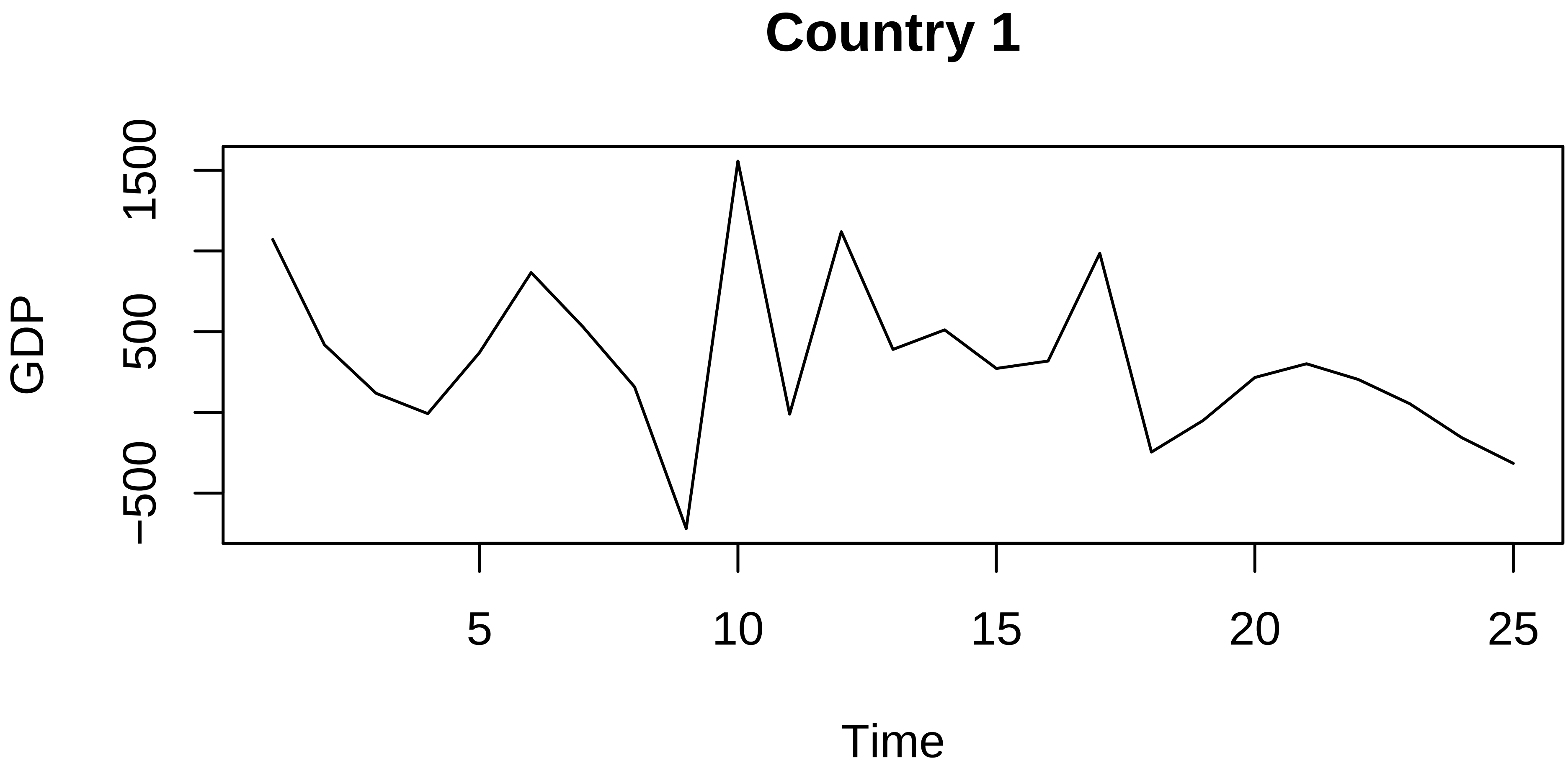
What do we think?

Clearly some heterogeneity

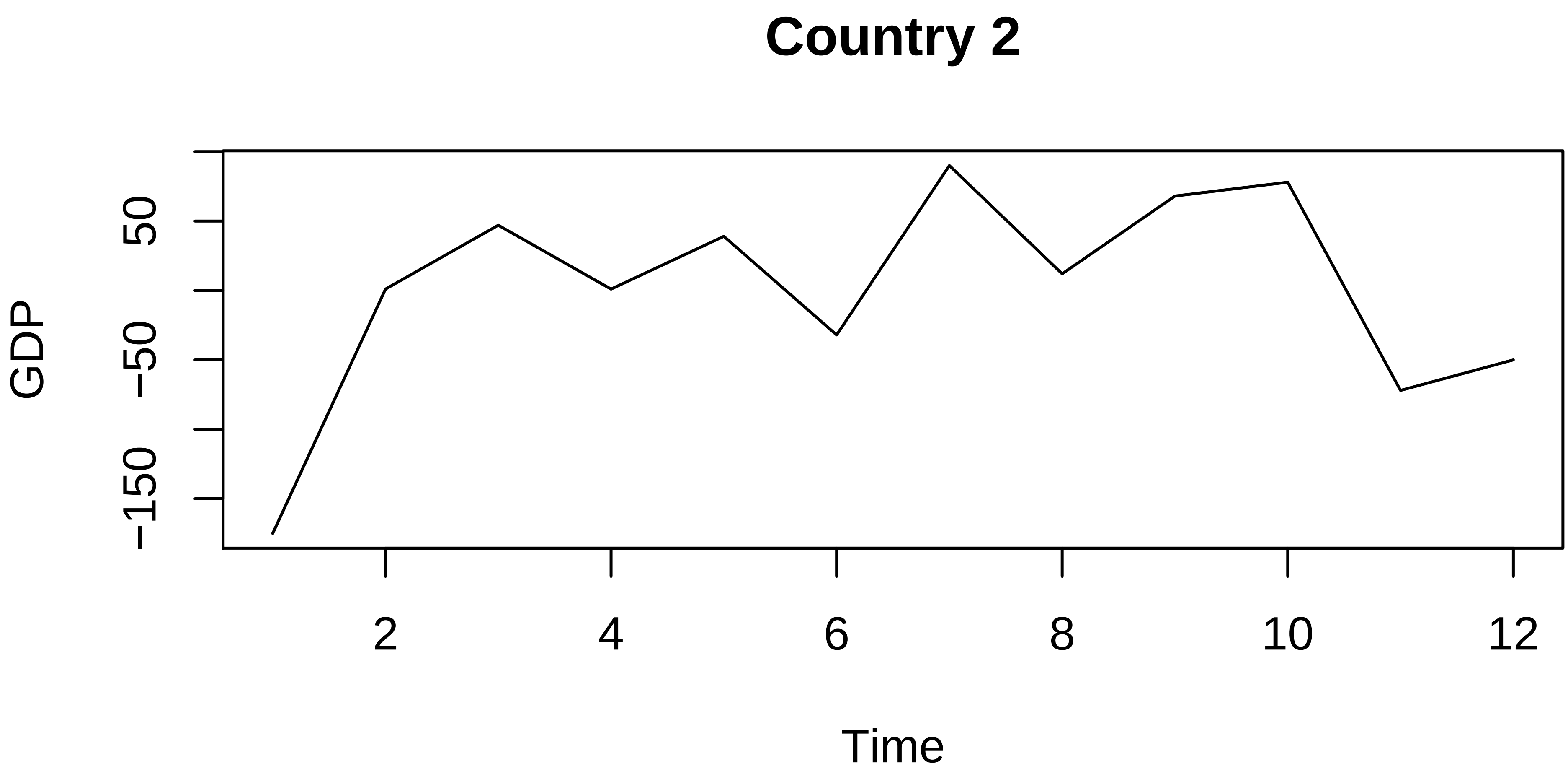
If had to pick one time series specification, choose ARIMA(0,1,0) or ARIMA(1,1,0)

Seems to fit many cases; guards against spurious regression

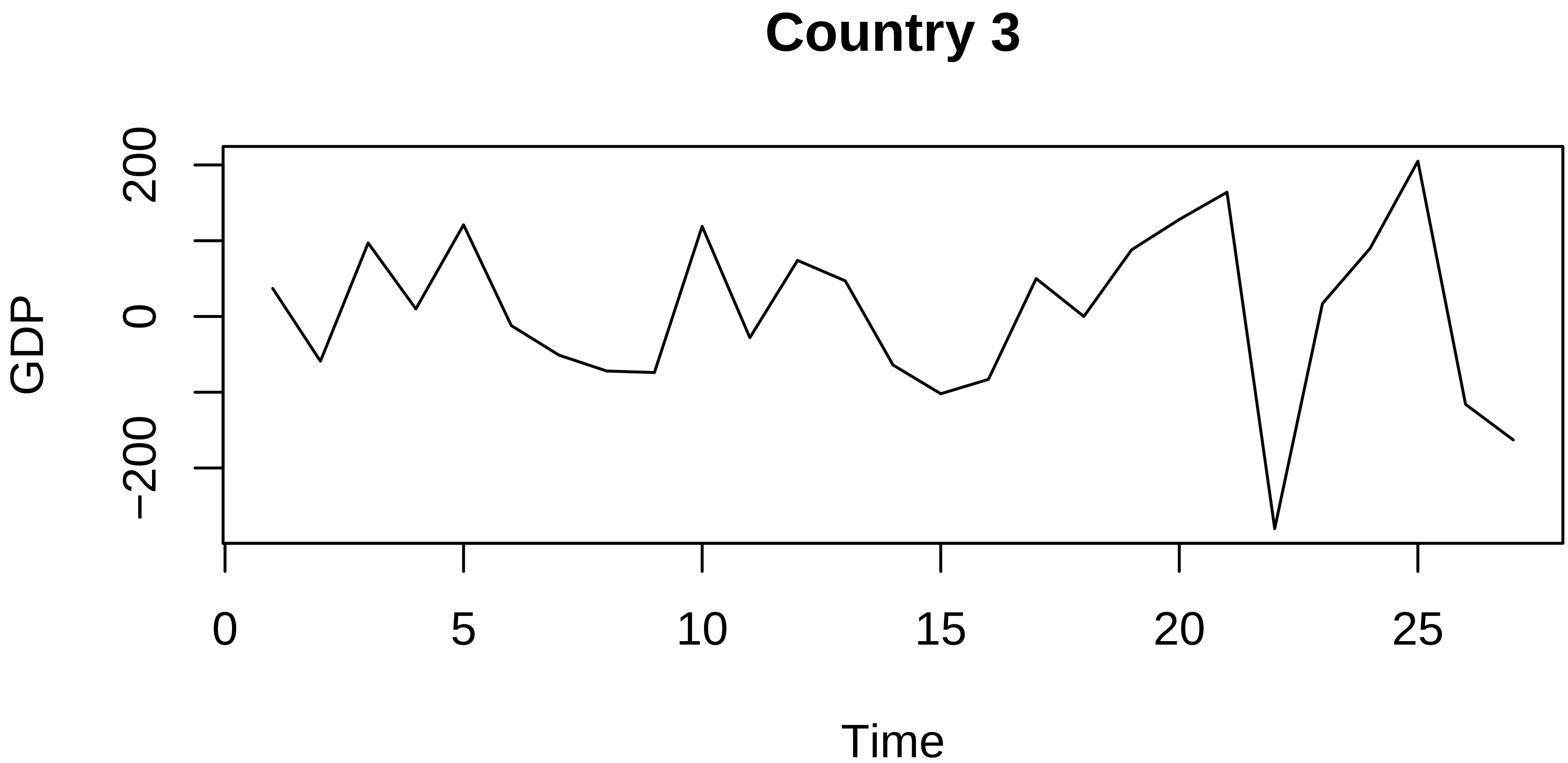
But if we're dubious about imposing a single ARIMA(p,d,q) across units,
we could let them be heterogeneous



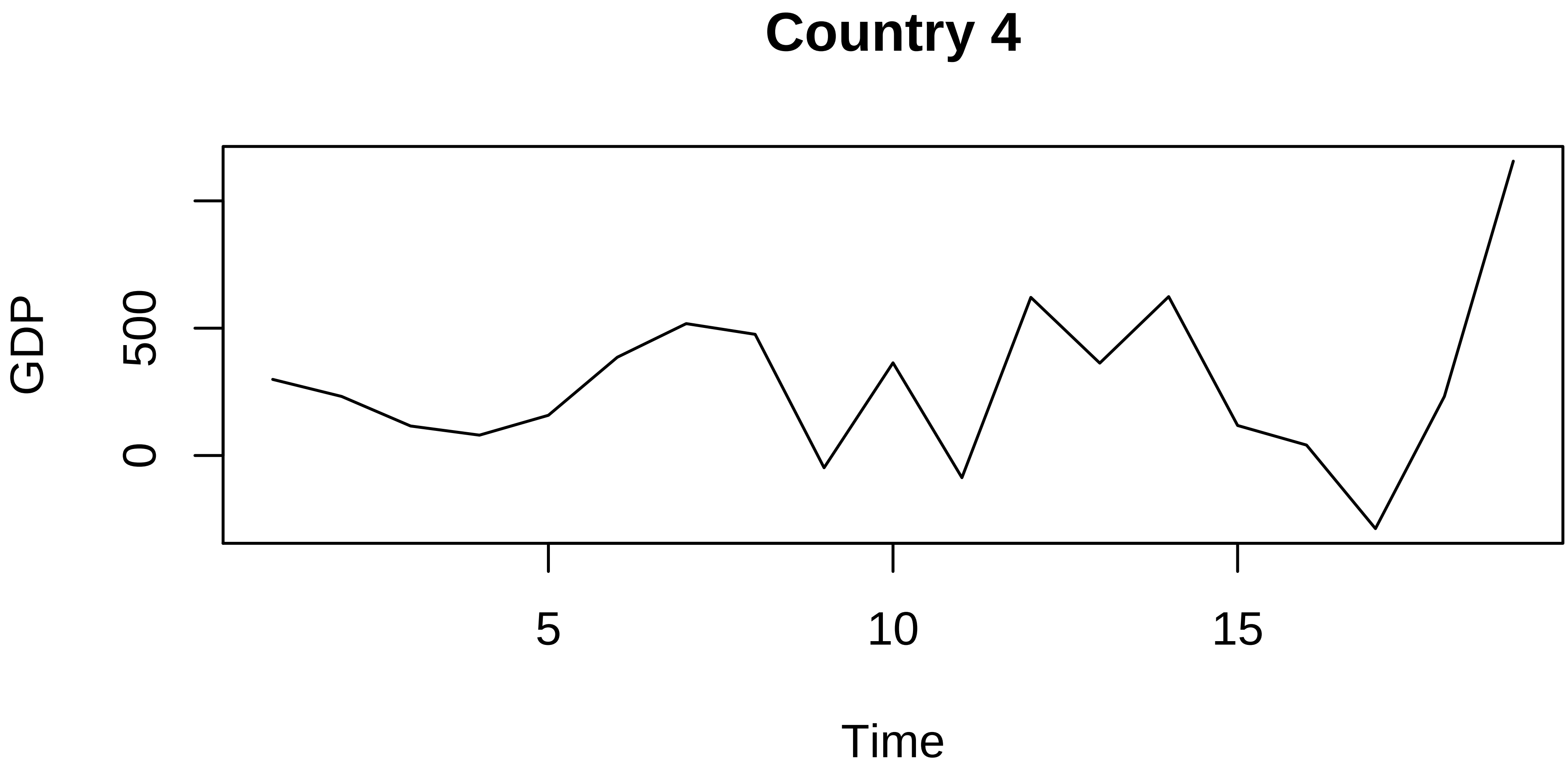
GDPdiff time series for country 1



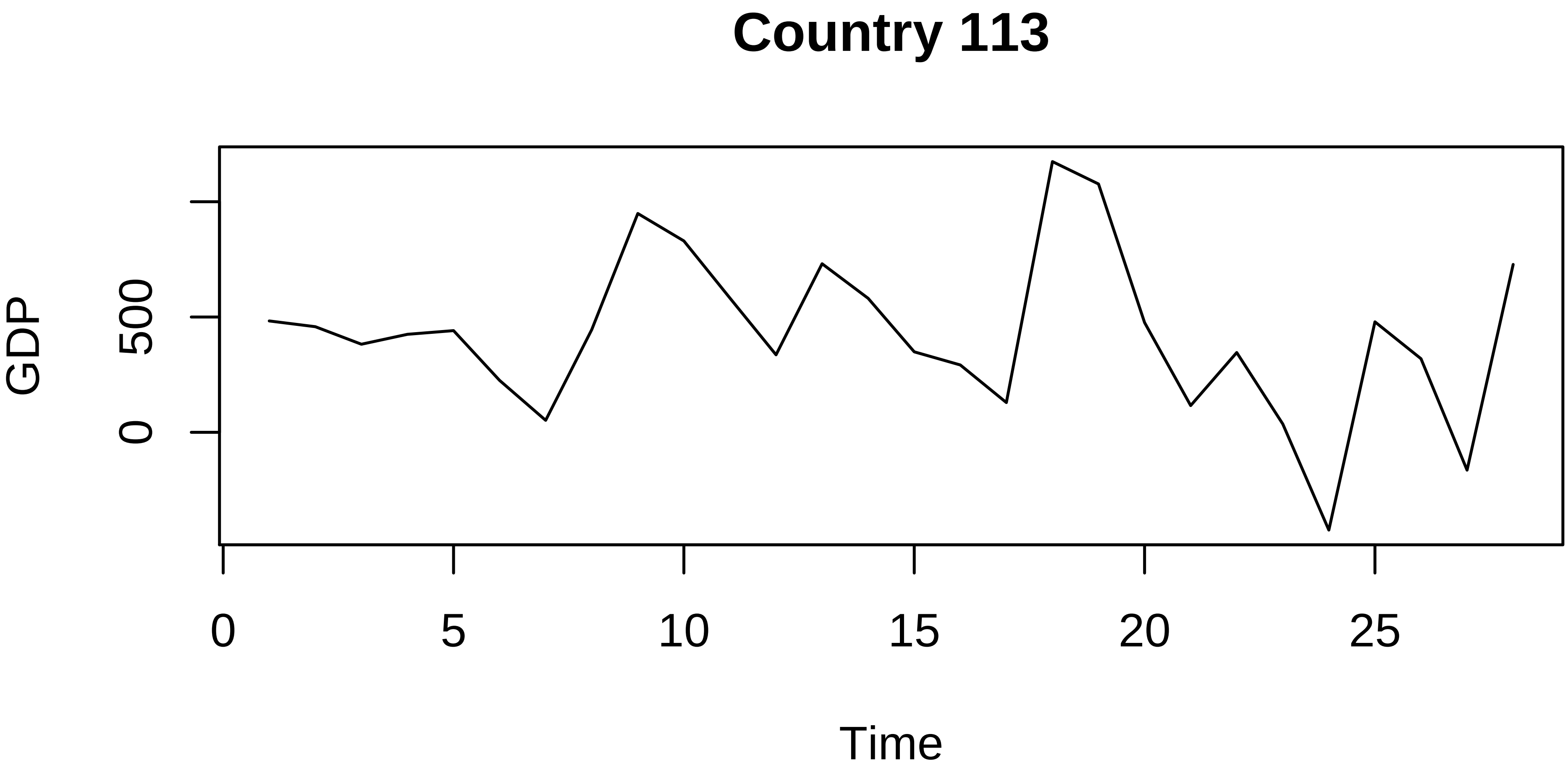
GDPdiff time series for country 2



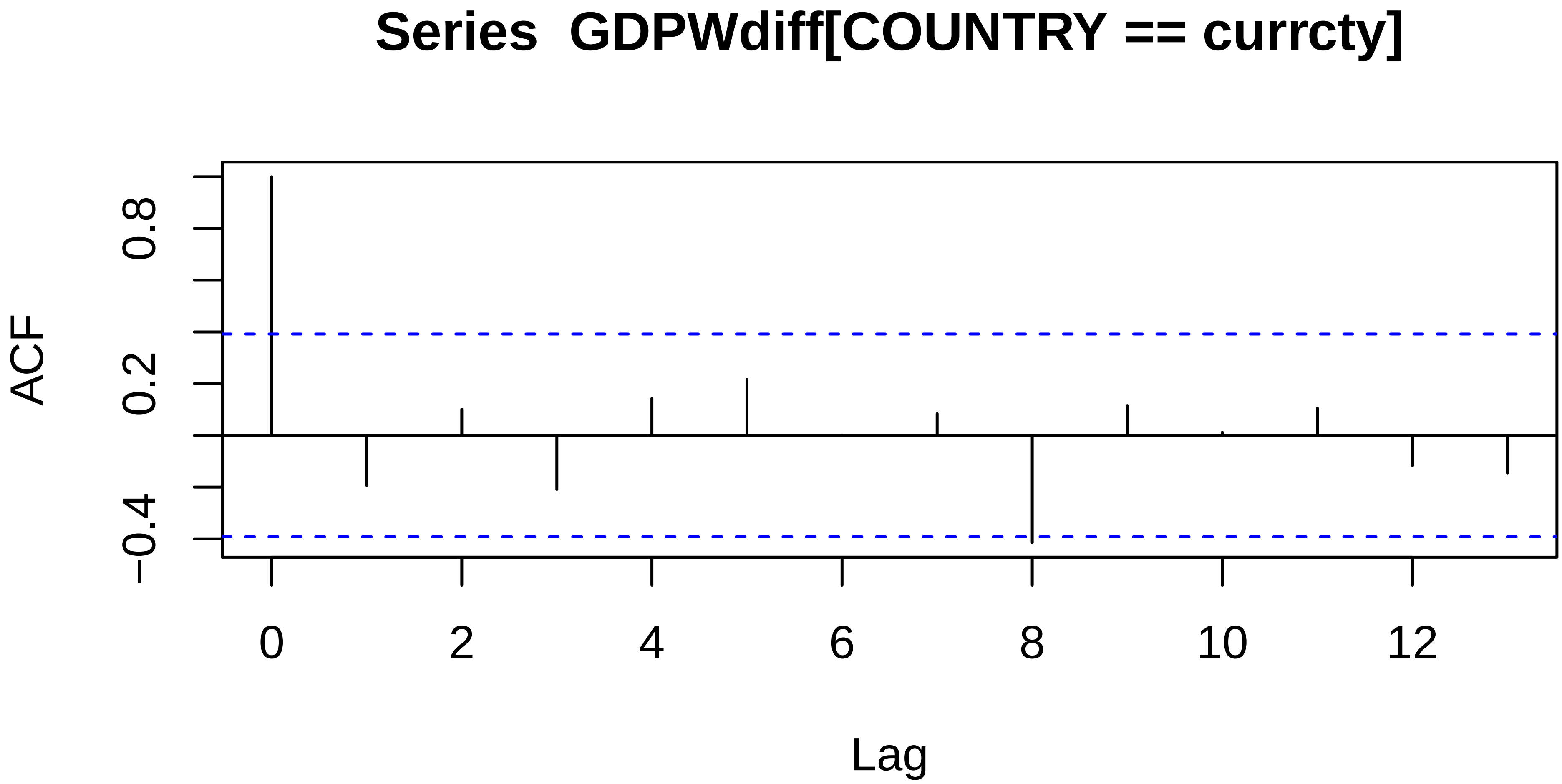
GDPdiff time series for country 3



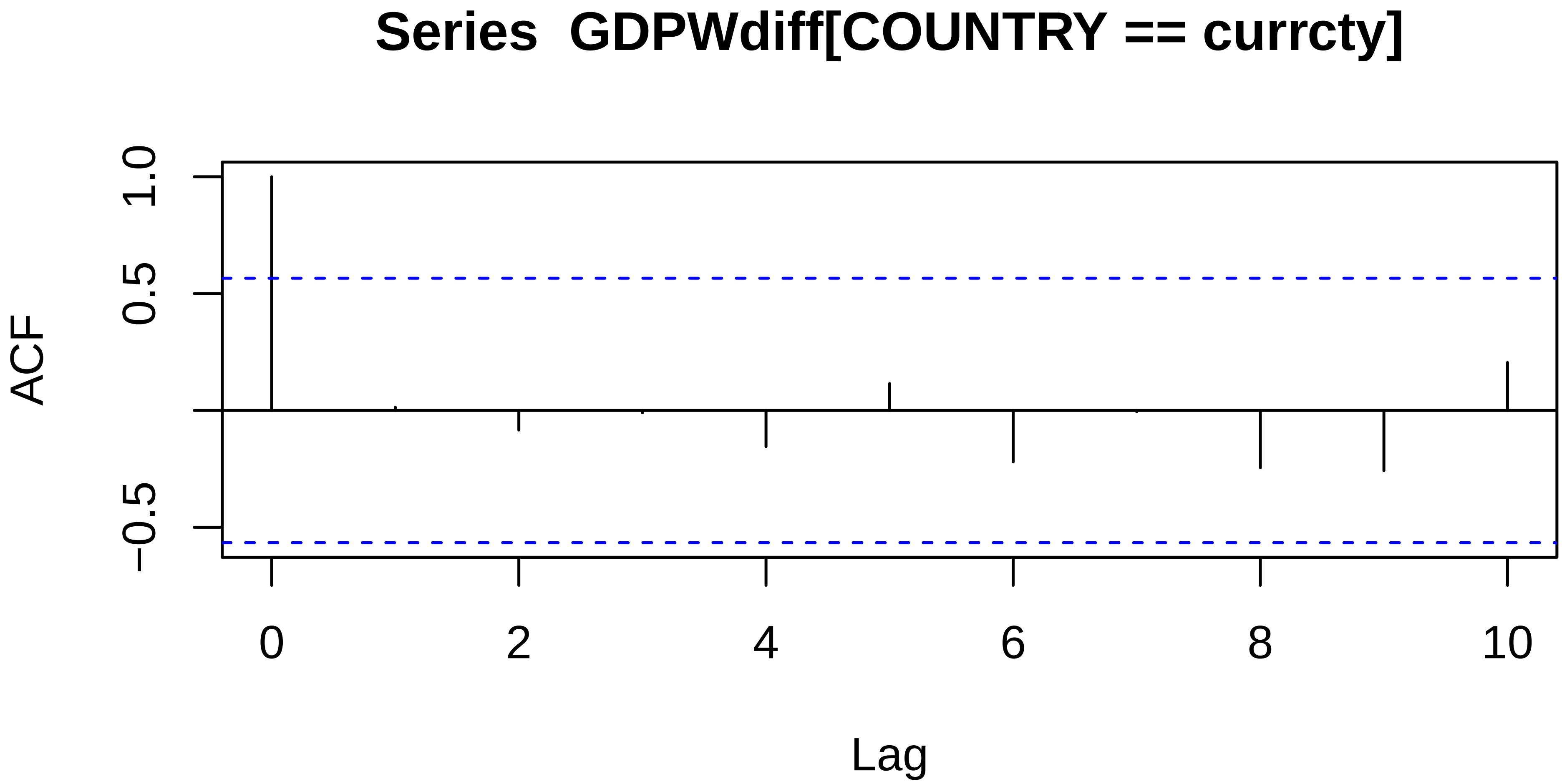
GDPdiff time series for country 4



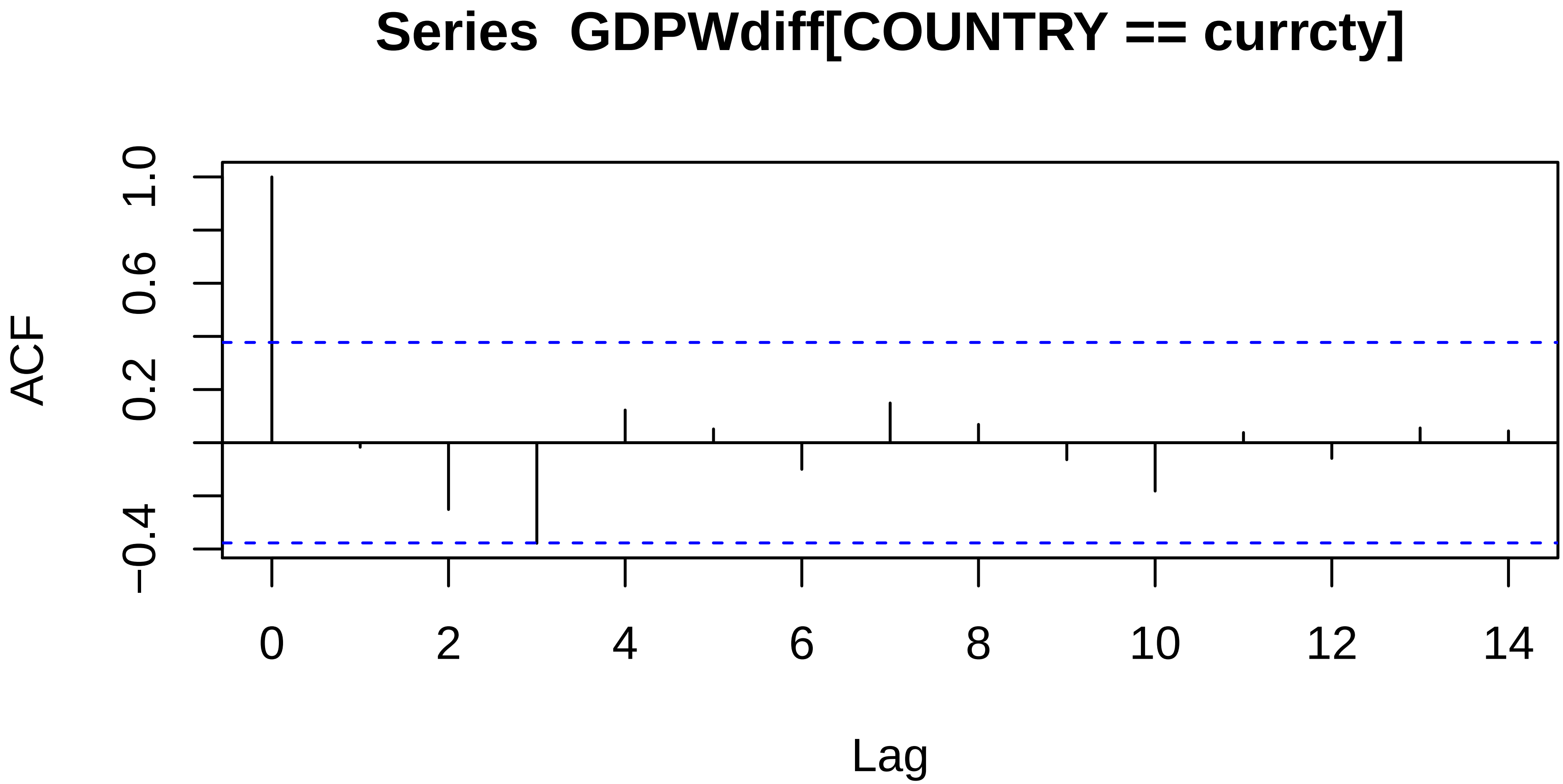
GDP diff time series for country 113



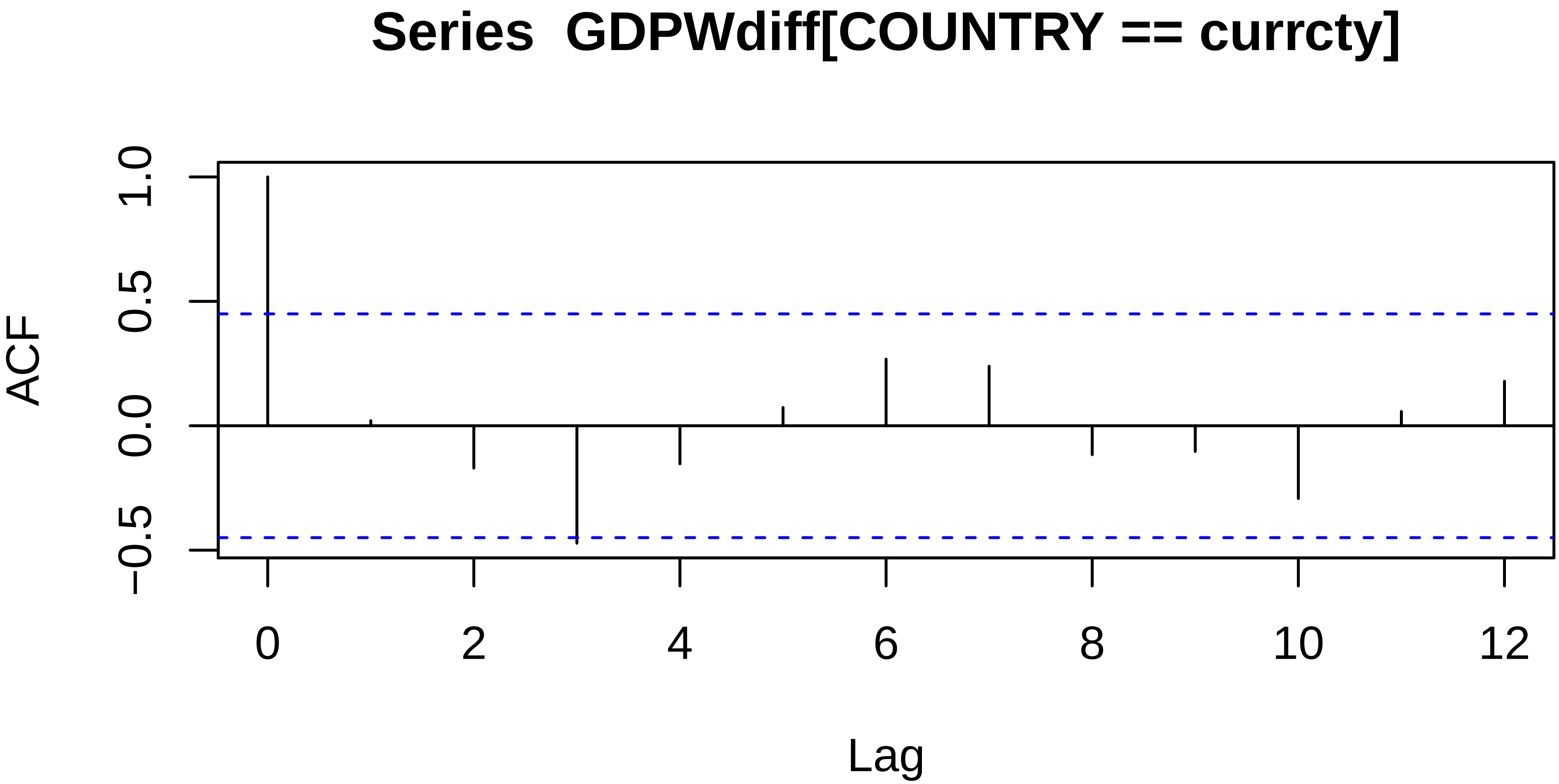
GDPdiff ACF for country 1



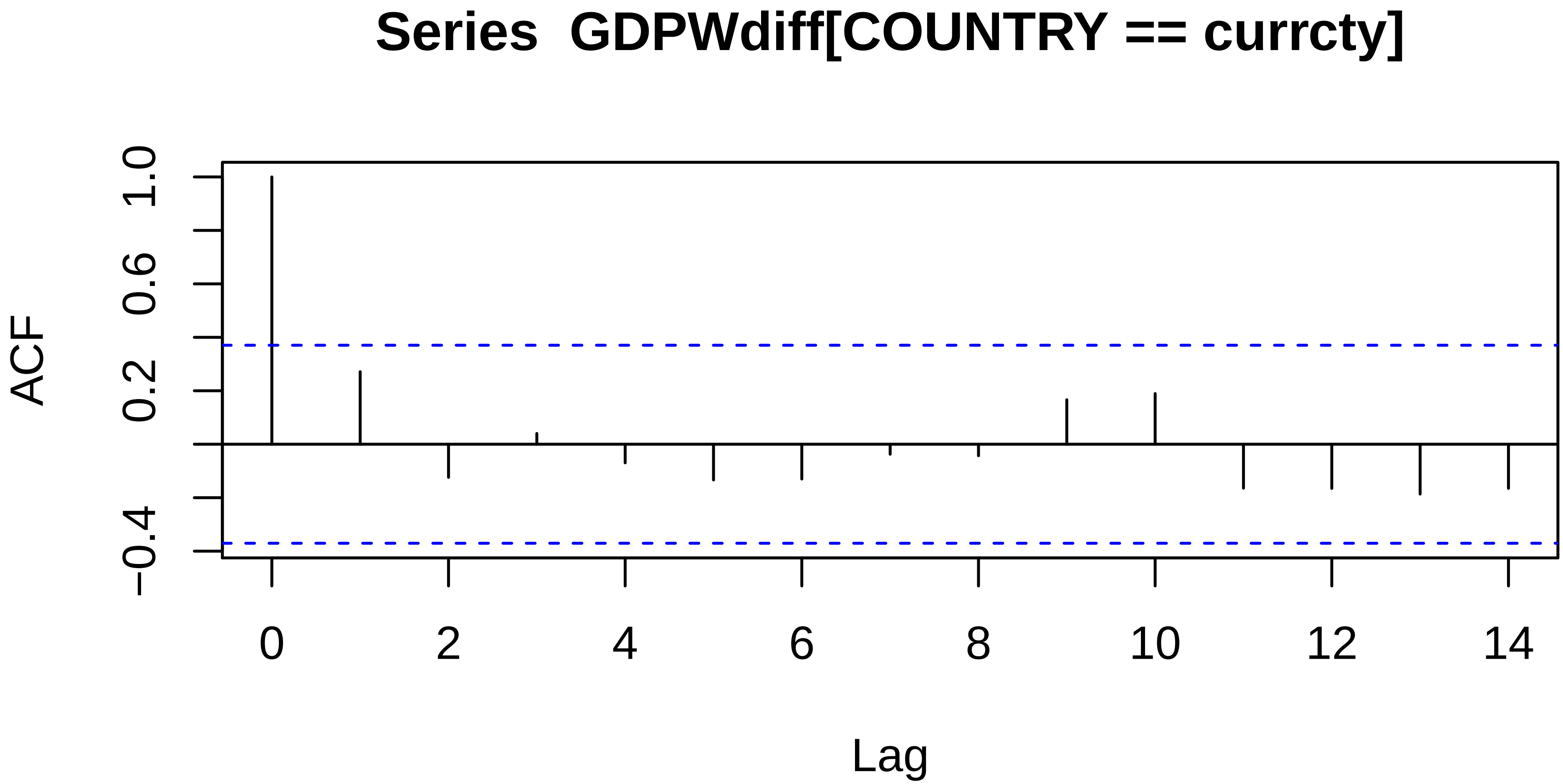
GDPdiff ACF for country 2



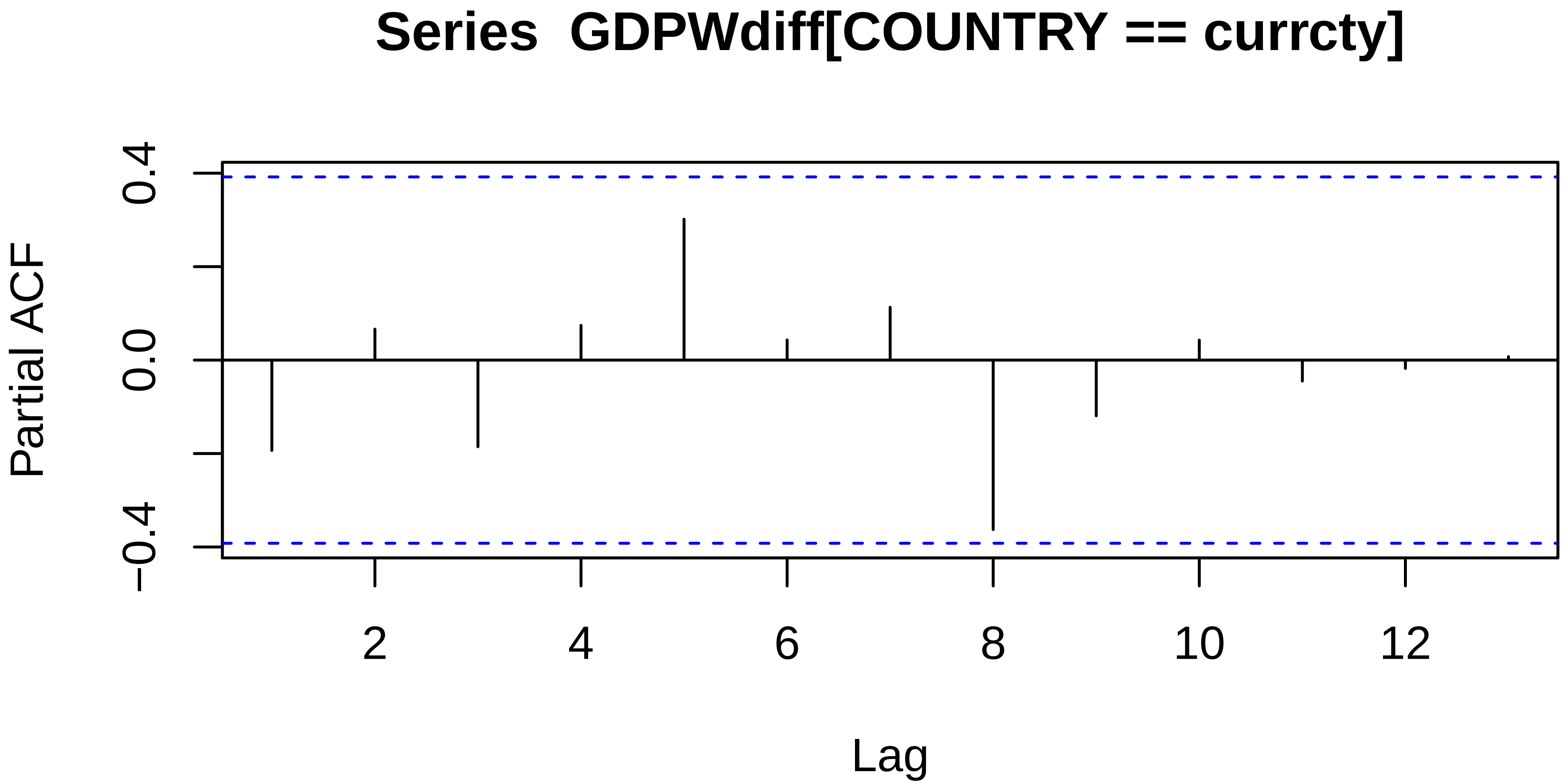
GDPdiff ACF for country 3



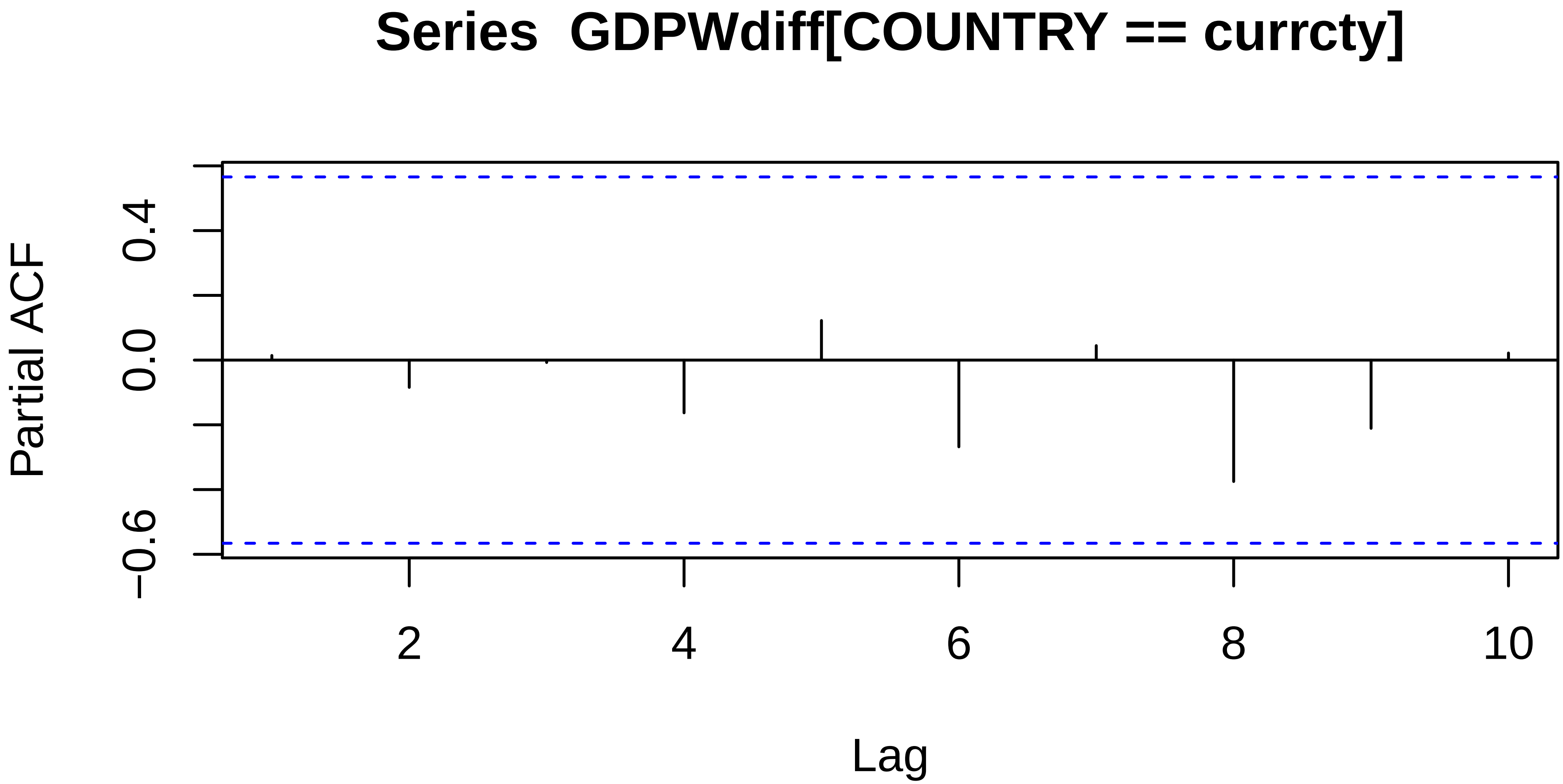
GDPdiff ACF for country 4



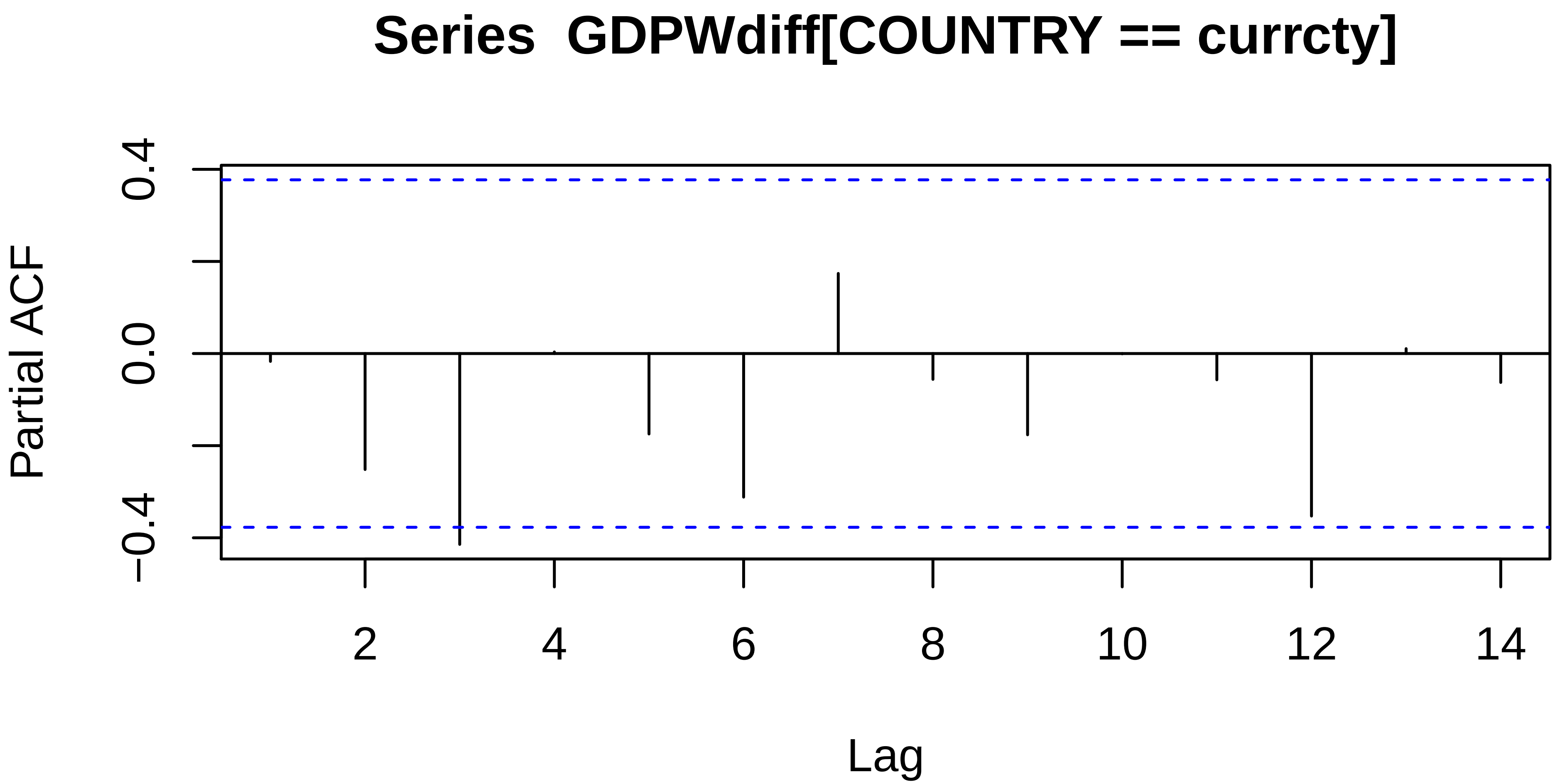
GDPdiff ACF for country 113



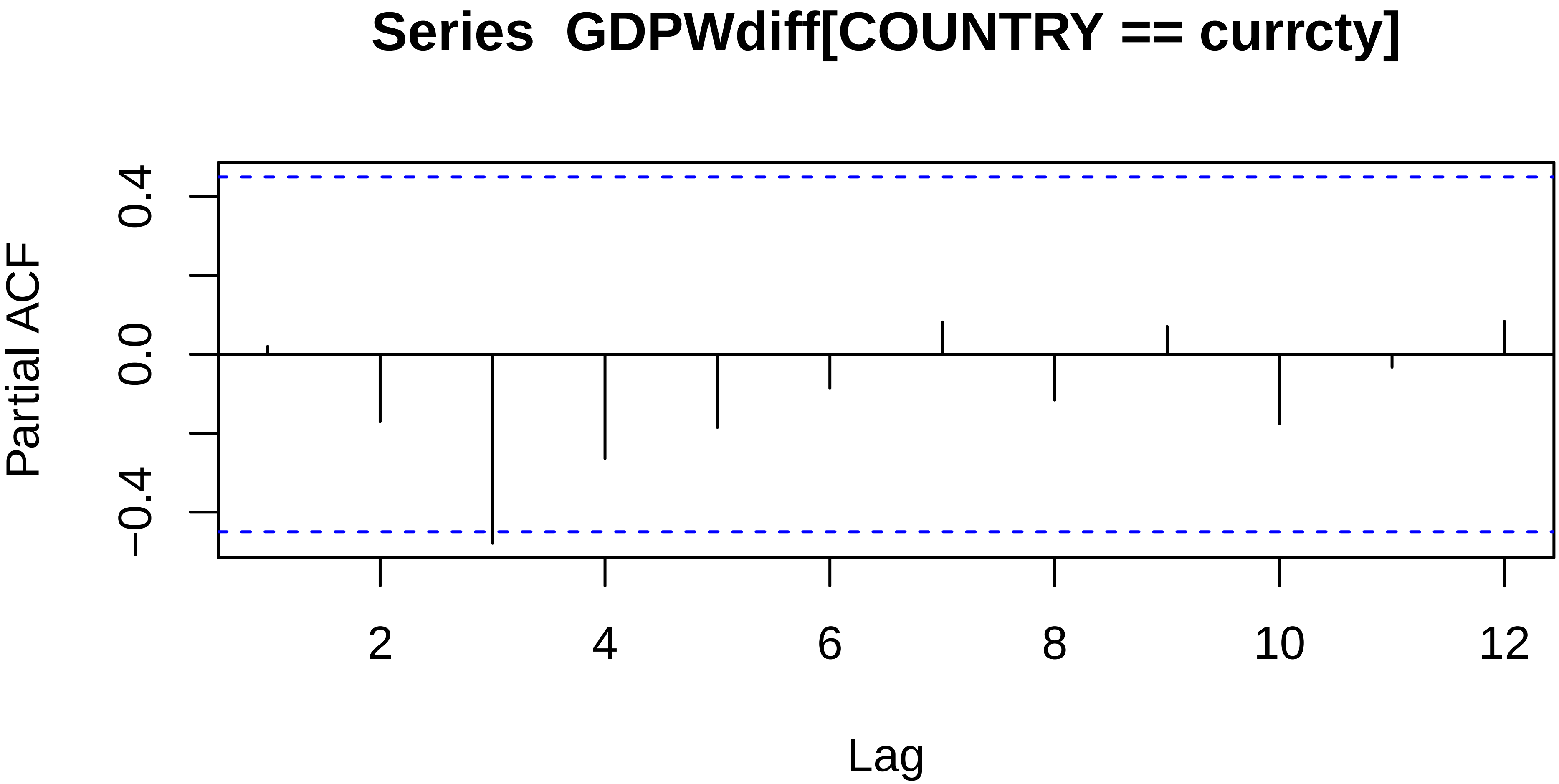
GDPdiff PACF for country 1



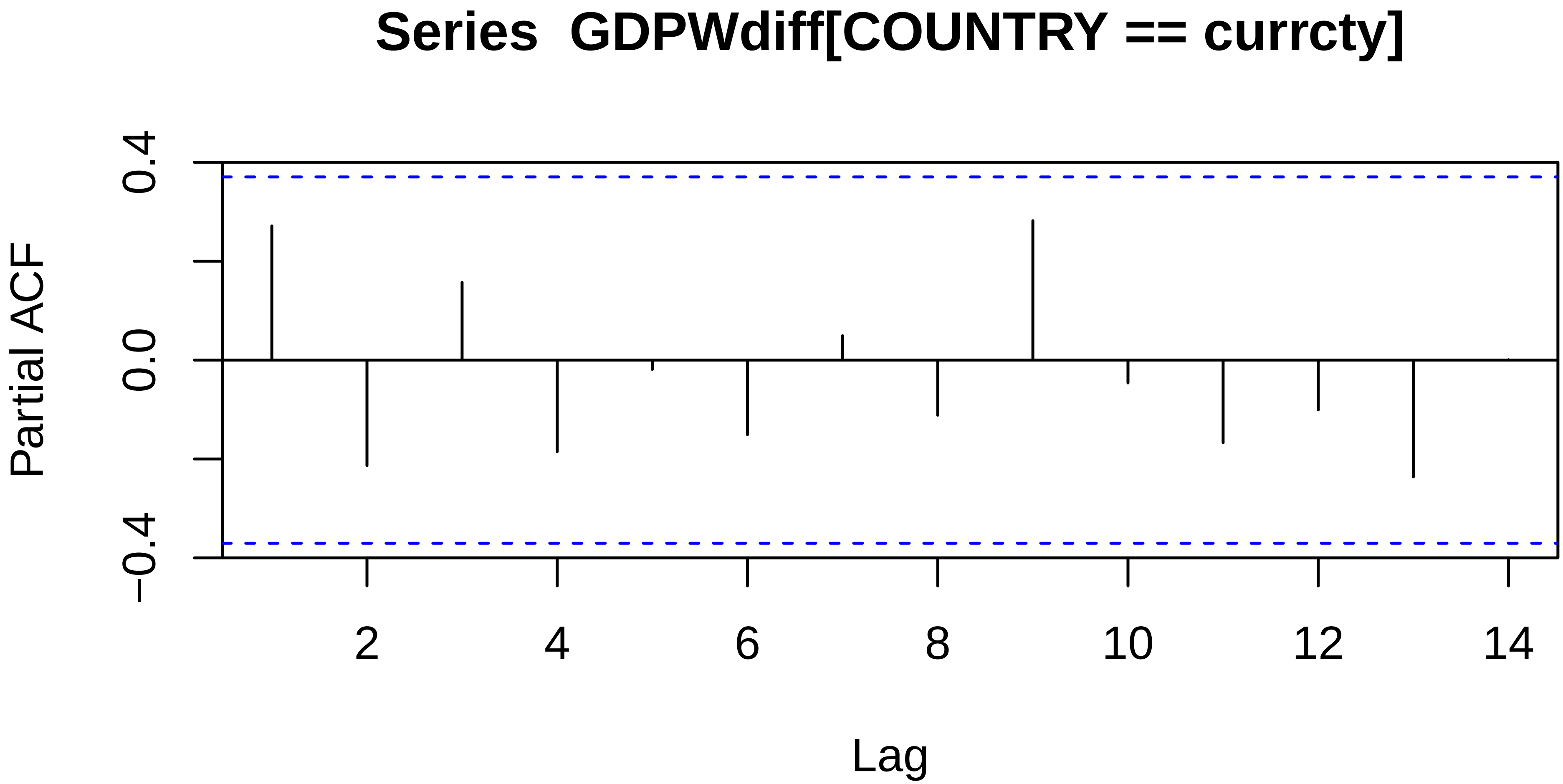
GDPdiff PACF for country 2



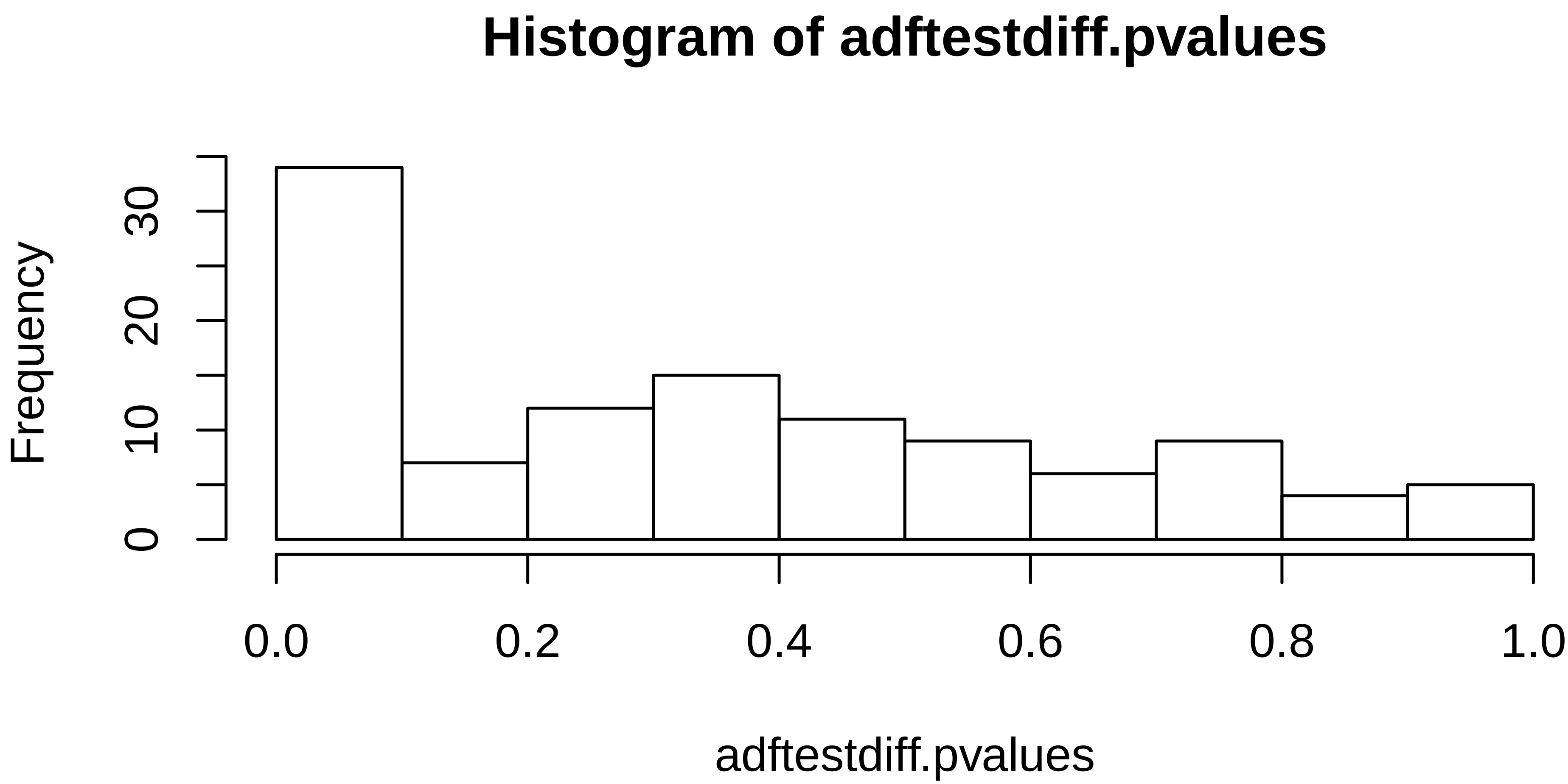
GDPdiff PACF for country 3



GDPdiff PACF for country 4

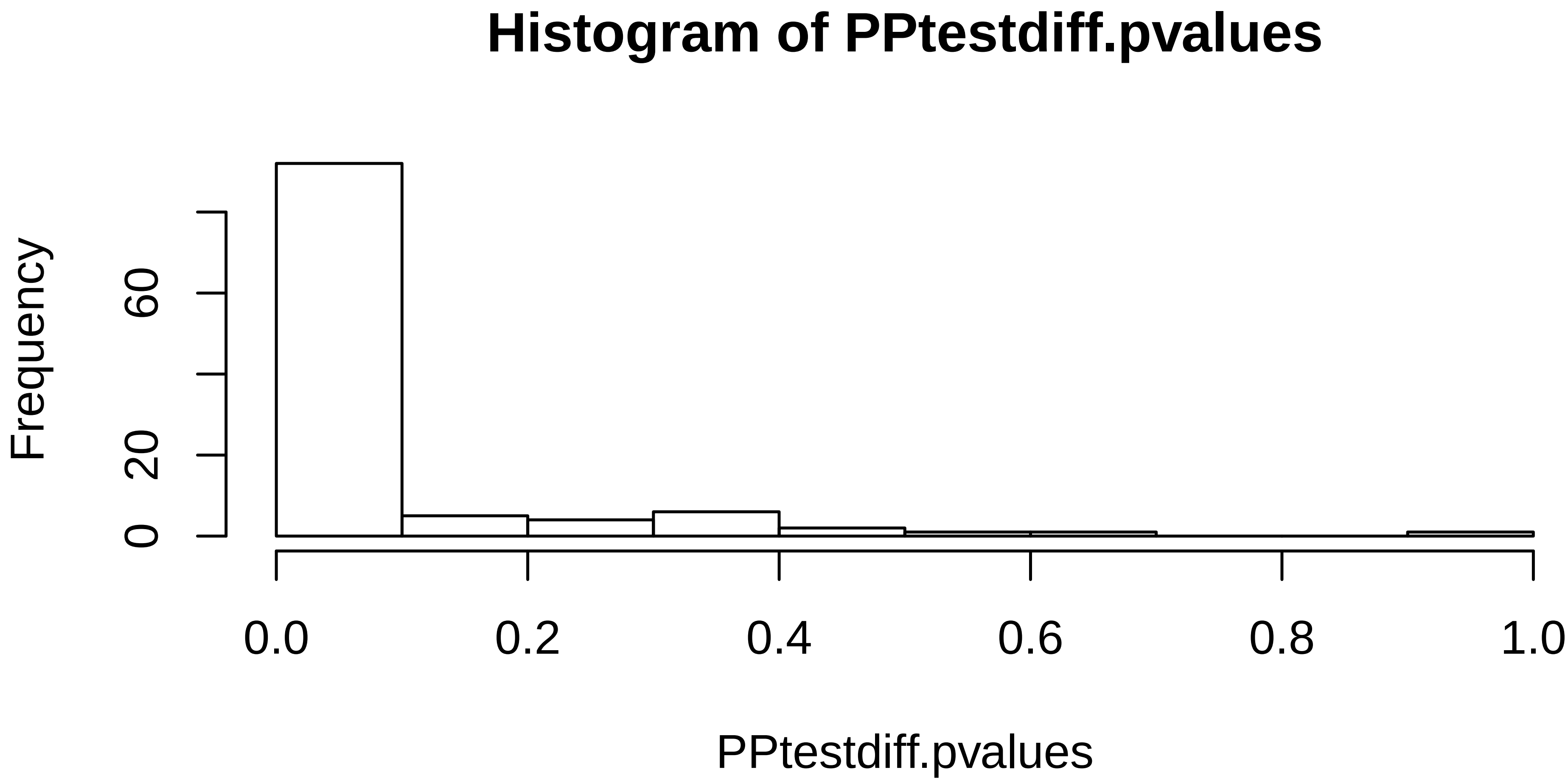


GDPdiff PACF for country 113



Histogram of p-values from ADF tests on GDPWdiff

What is this pattern consistent with?



Histogram of p-values from Phillips-Peron tests on GDPWdiff

Example continued in R demonstration

We will continue this example in section using the code provided

For now, let's focus on the results that emerge,
and how they depend on treating intercepts as either random or fixed by country

In particular, we want to see if fixed effects can help us with omitted time invariant variables, which are legion in this example

In the example, we will decide on an ARIMA(1,1,0) model of GDP
(What does this mean?)

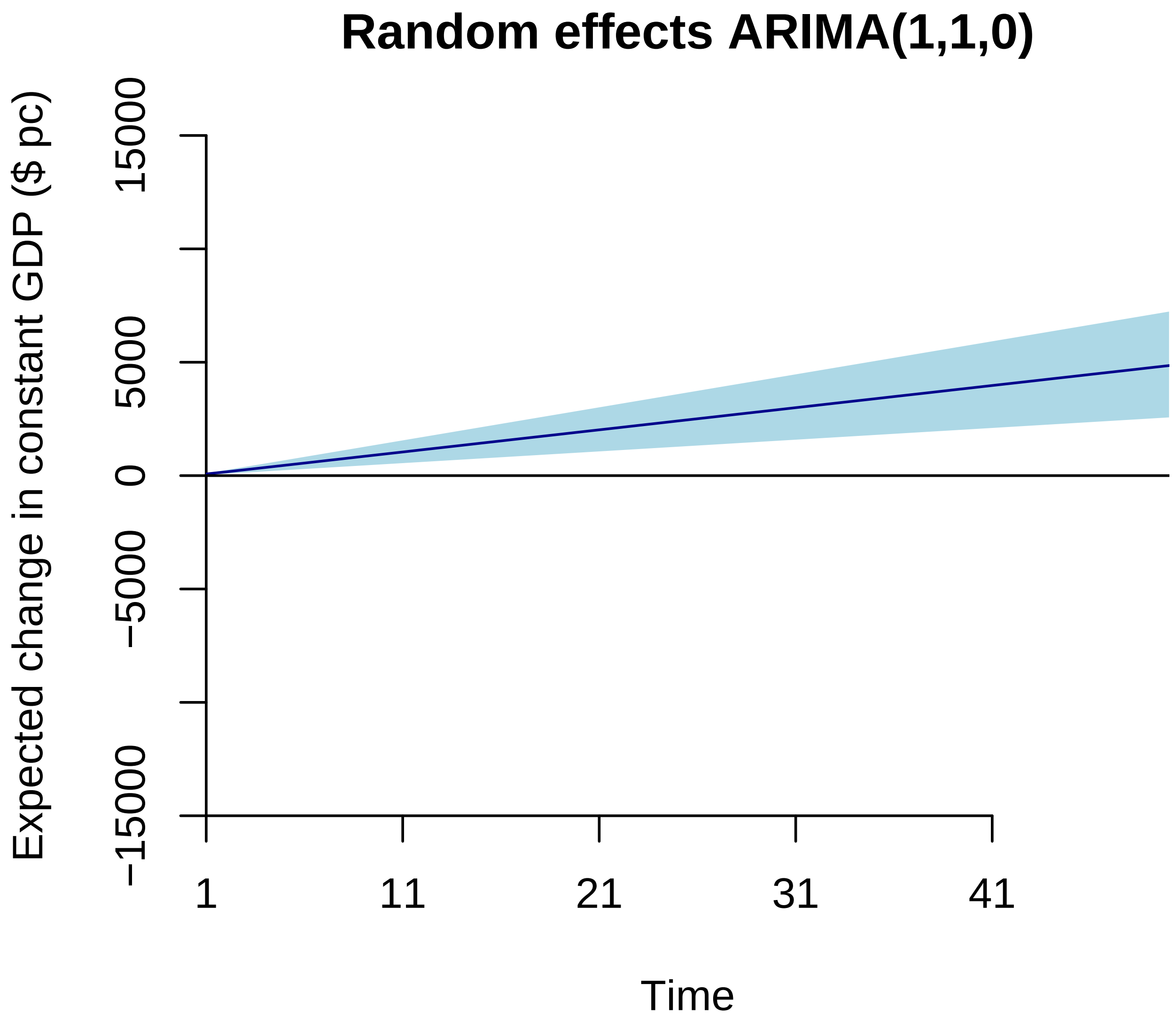
We will fit three different models of the relationship between education and GDP:

1. ARIMA(1,1,0) with random country intercepts
and controls for oil producing countries and democracy

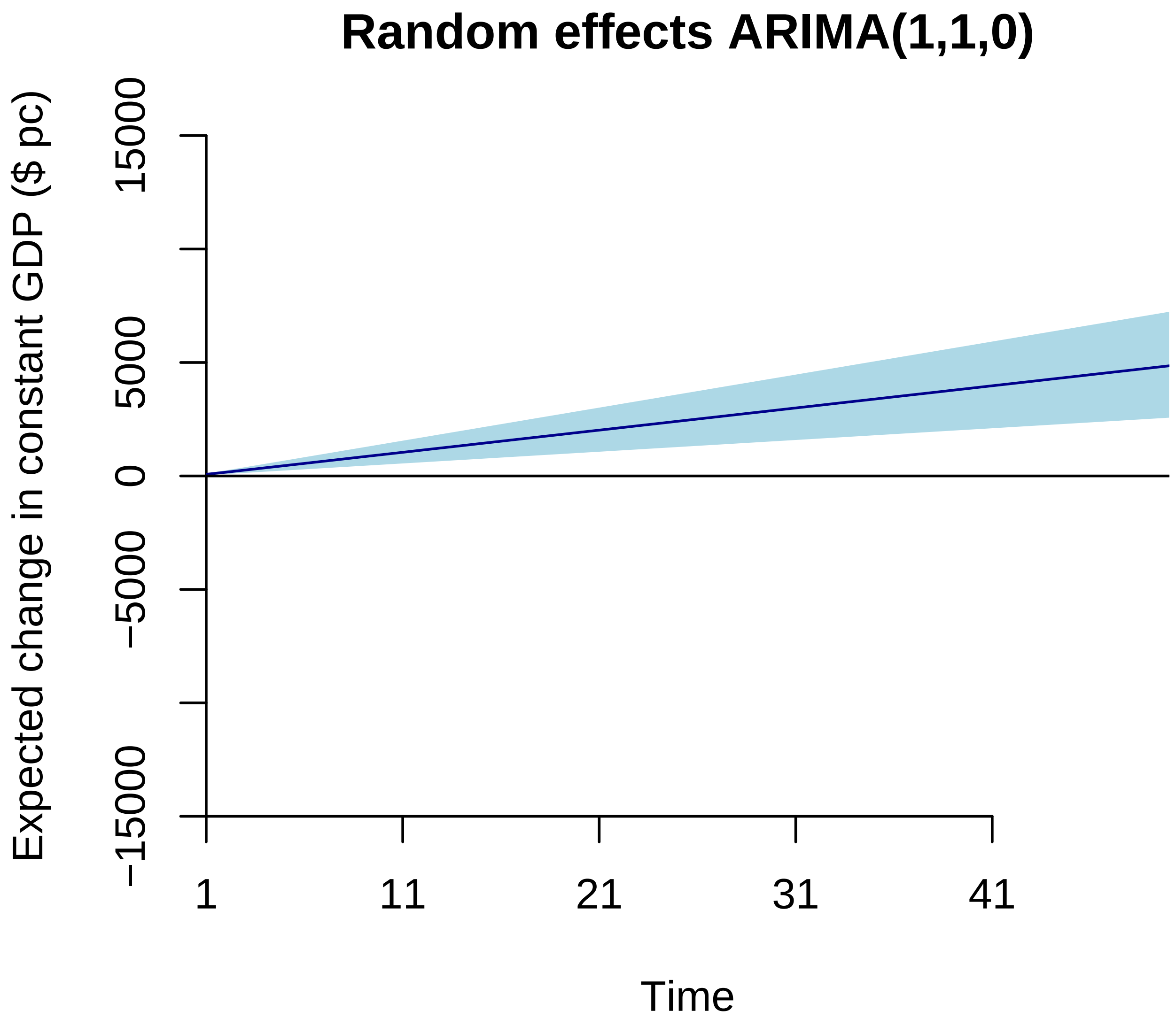
2. ARIMA(1,1,0) with fixed country intercepts
and controls for democracy

3. ARIMA(1,1,0) with "mixed" country intercepts
and controls for democracy

	Model				
	RE	FE	FE-pcse	FE	"ME"
Education$_{it}$	110.10	−75.56	−75.56	−86.68	−84.75
	33.99	12.16	13.48	12.45	14.56
Democracy$_{it}$	−20.25	−12.90	−12.90	−26.15	−3.63
	43.83	47.69	50.58	47.97	55.22
Oil-Producer$_{it}$	28.05	—	—	—	—
	27.51	—	—	—	—
GDP$_{i,t-1}$	0.23	0.15	0.15	0.17	0.20
		0.02	0.02	0.02	
GDP$_{i,t-2}$				−0.12	
				0.02	
σ_α	0.14	—	—	—	309.10
Fixed effects?		x	x	x	x
Random effects?	x				x
N	113	113	113	113	113
T	3–28	3–28	3–28	2–28	328
obs. $N \times T$	2794	2794	2794	2741	2794
AIC	43376				42113
LM test p-value		0.001	0.001	0.131	

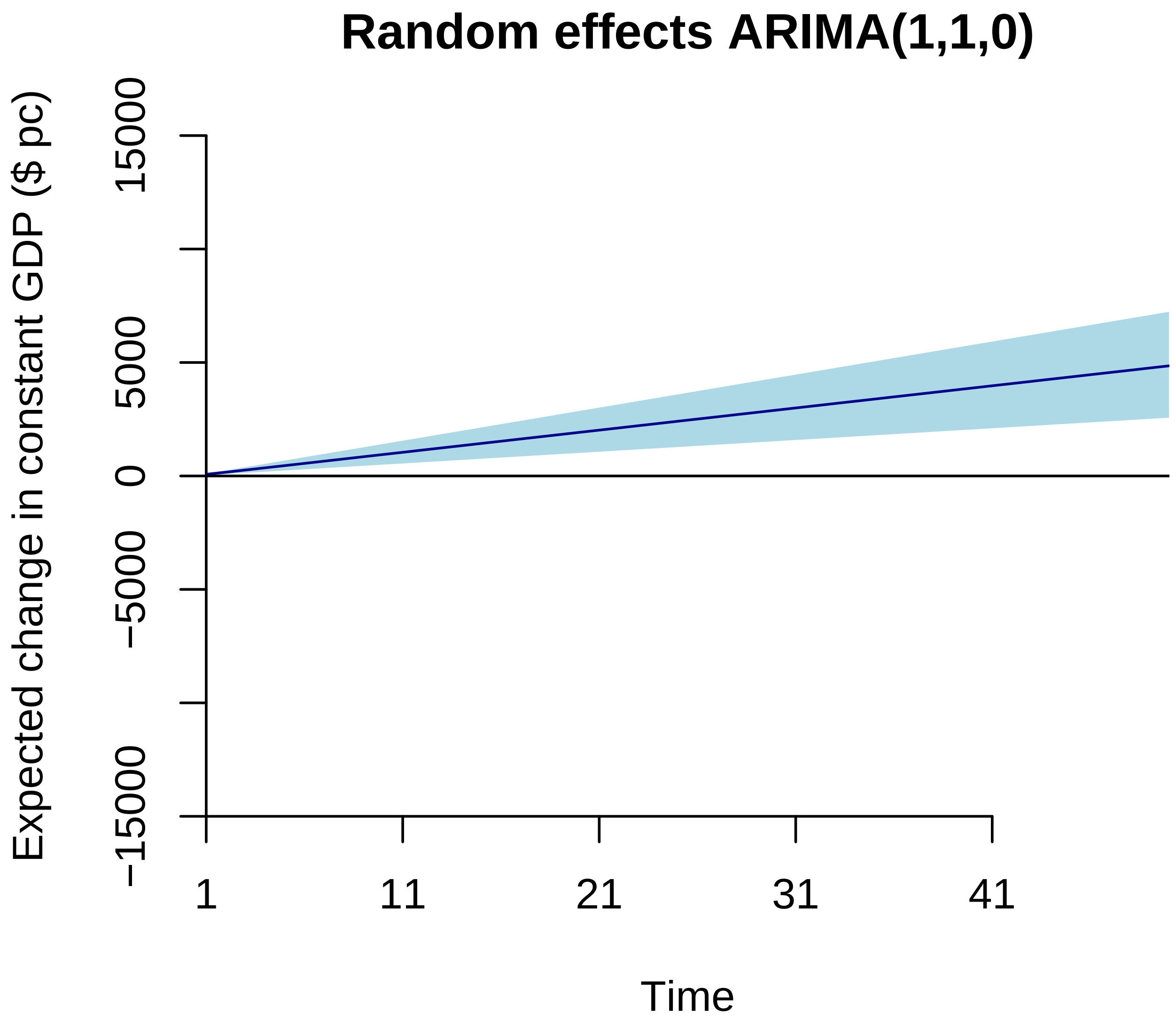


What does the model imply substantively,
and how does this depend on model assumptions?

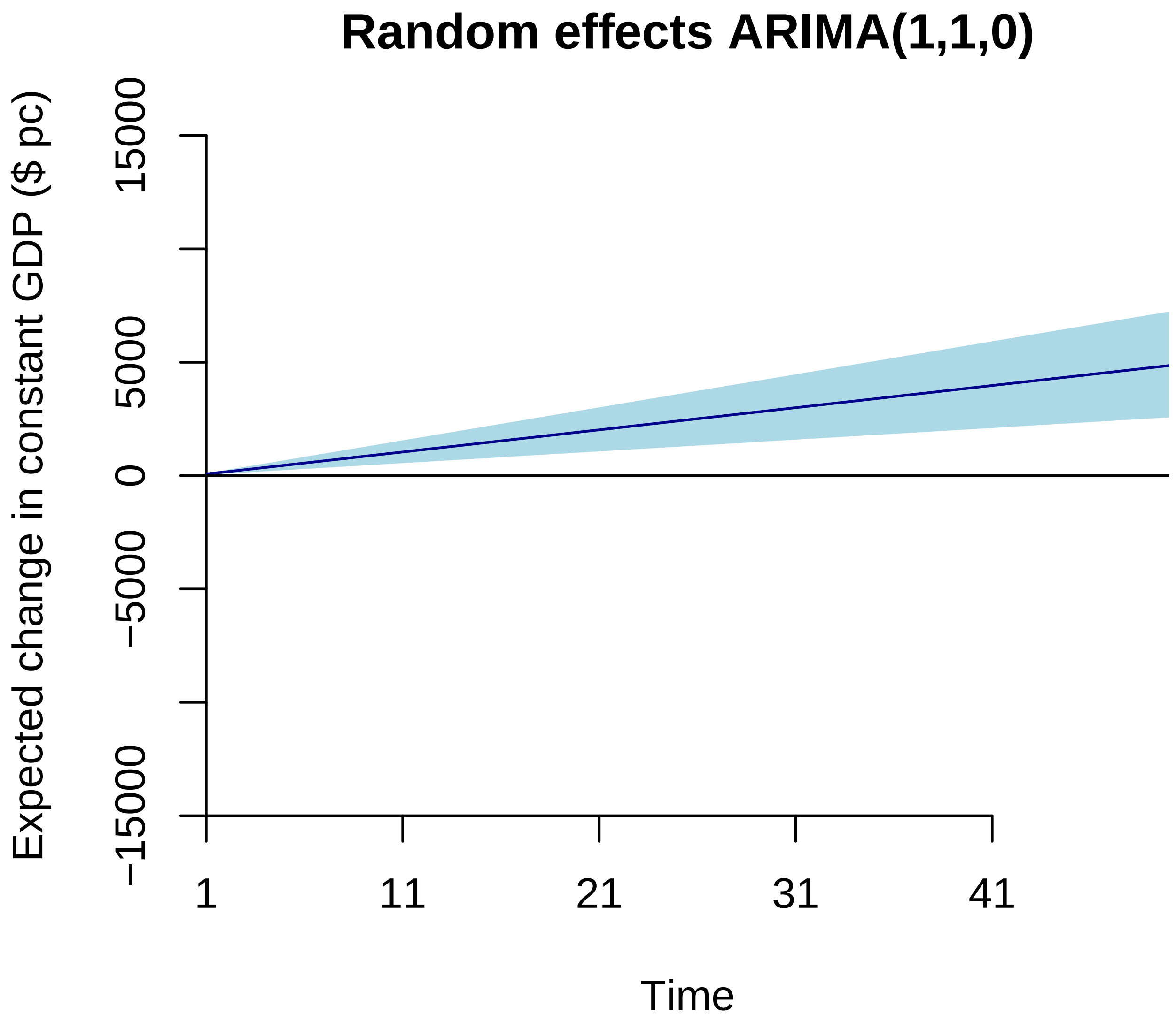


Suppose we have a country with "average" characteristics,
and we increase education to 1 sd above the mean

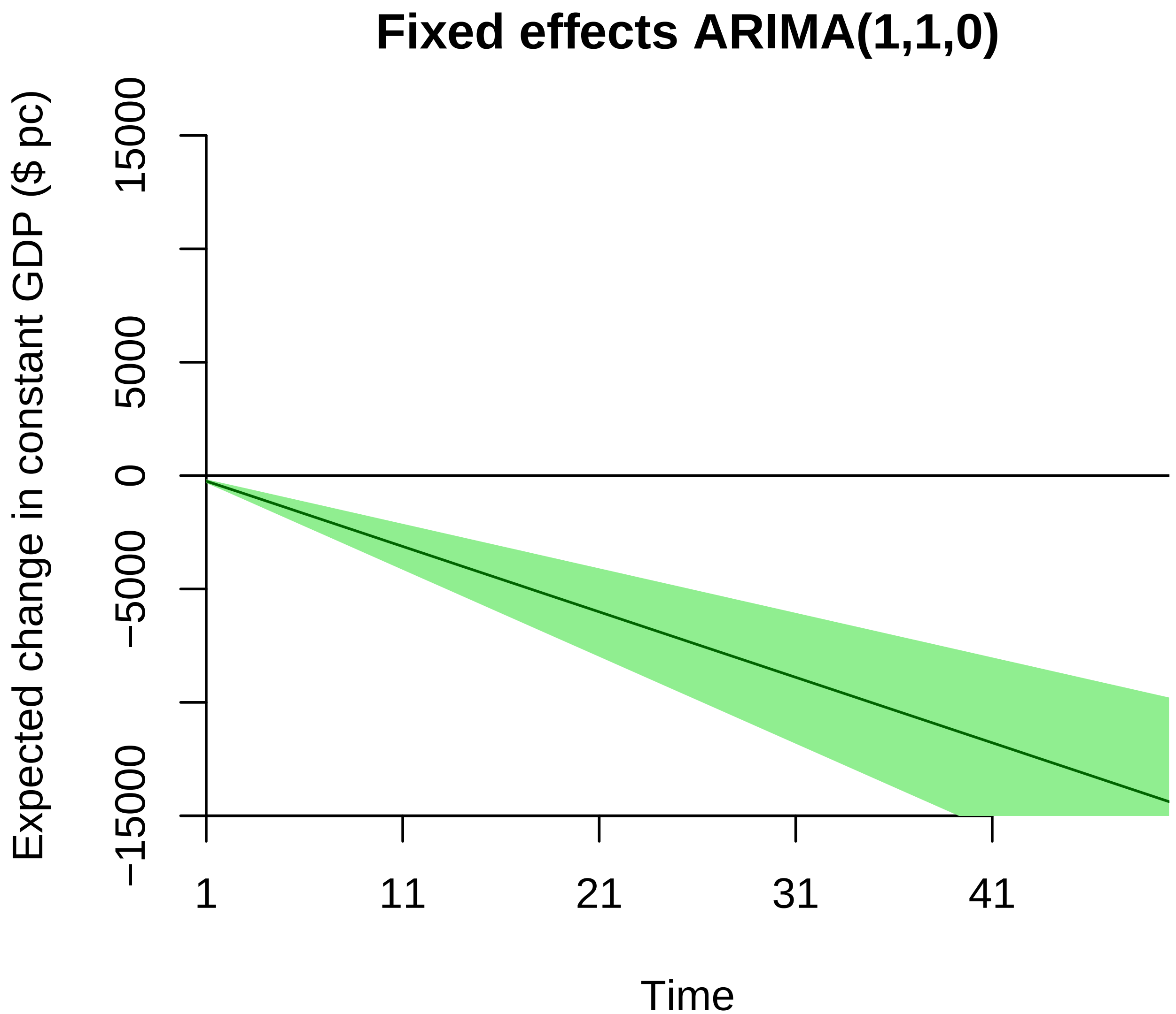
How much does the model predict education to rise over the following years?



The above plot shows the expected change in GDP over time in the high education country relative to an average (untreated) country

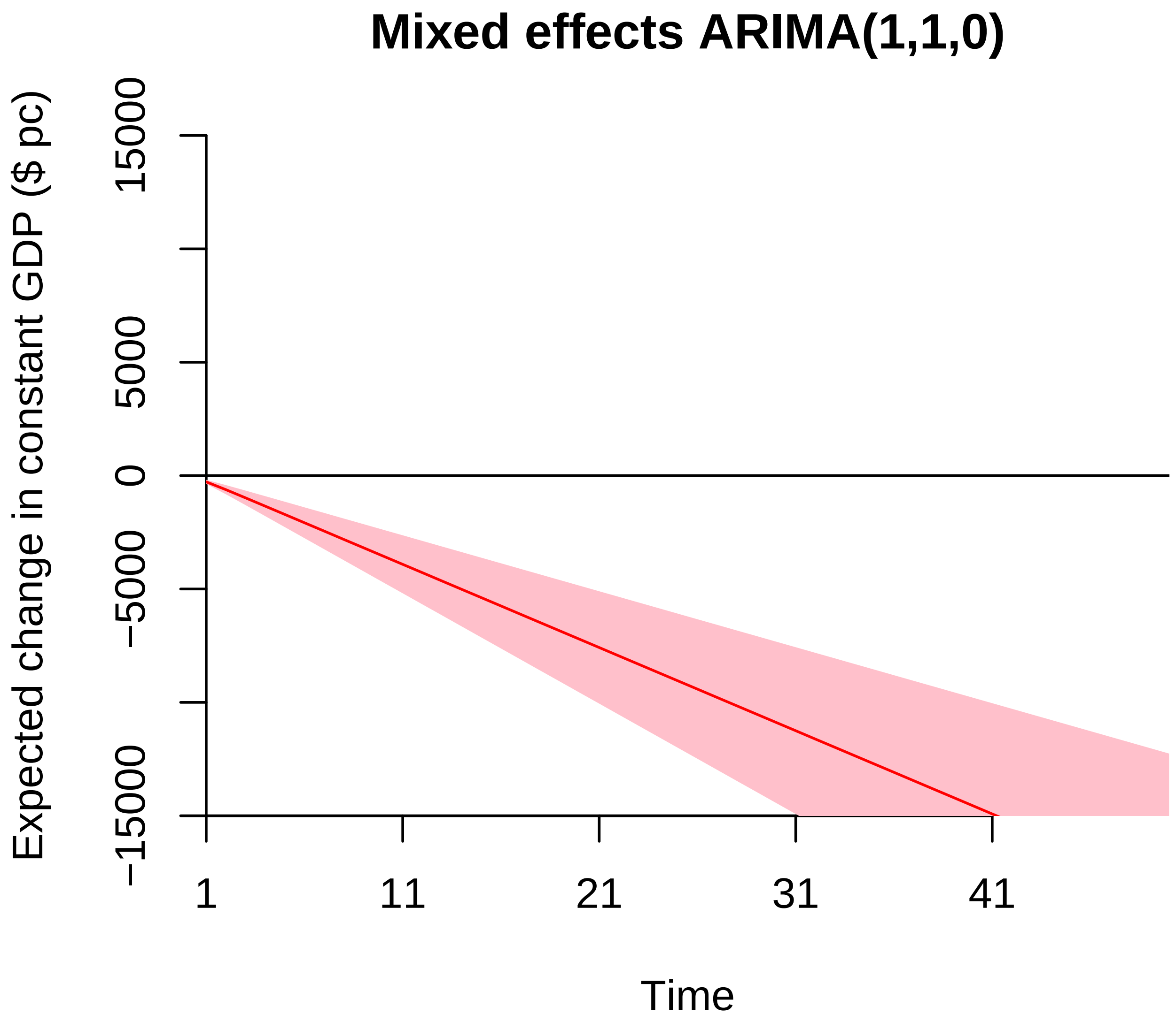


The result seems sensible, but the model:
(1) ignored many unmeasured confounders and (2) differences GDP,
so we should be skeptical in both the short- and long-run



Adding fixed effects completely flips the results for education

Now the results make little sense
(Suggests the model is badly identified, even with fixed effects)



In a model with both RE and FE for countries, the FEs dominate, as the "Mixed" effects model shows