

Climate Effects on GDP Growth (Kahn et al., 2021, Energy Economics)

Samuele Centorrio scentorrio@proton.me

2025-05-02

Methodology in Kahn et al. (2021)

Kahn et al. (2021) start from the following ARDL specification (see Kahn et al. (2021), p. 4 eq 1)

$$\Delta y_{i,t} = a_i + \sum_{\ell=1}^p \varphi_{\ell} \Delta y_{i,t-\ell} + \sum_{\ell=0}^p \beta_{\ell} \Delta \tilde{x}_{i,t-\ell}(m) + \varepsilon_{i,t}, \tag{1}$$

where i is a country, t is time, $y_{i,t}$ is the logarithm of real GDP for country i and time s , $\tilde{x}_{i,t}$ is a vector of regressors which include climate variables such as temperature and precipitations, taken in deviations from their historical averages¹, a_i is a country fixed effect, and finally Δ is used to denote first-differences. $\Delta y_{i,t}$ is thus equal to the growth rate of country i at time t which, for simplicity, we denote $r_{i,t}$. Thus, we rewrite

$$r_{i,t} = a_i + \sum_{\ell=1}^p \varphi_{\ell} r_{i,t-\ell} + \sum_{\ell=0}^p \beta_{\ell} \Delta \tilde{x}_{i,t-\ell}(m) + \varepsilon_{i,t}. \tag{2}$$

The ARDL model in equation (2) can be estimated "as is". However, it is possible that the growth rate, $r_{i,t}$ and $\Delta \tilde{x}_{i,t-\ell}(m)$ are cointegrated (e.g., a long-run equilibrium relationships between these variables may exist), and therefore the authors of the paper estimate an **error correction version of this model**. That is,

$$\begin{aligned} r_{i,t} &= a_i + \sum_{\ell=1}^p \varphi_{\ell} r_{i,t-\ell} + \sum_{\ell=0}^p \Delta \tilde{x}_{i,t-\ell}(m) \beta_{\ell} + \varepsilon_{i,t} \\ \Delta r_{i,t} &= a_i - r_{i,t-1} + \sum_{\ell=1}^p \varphi_{\ell} r_{i,t-\ell} + \sum_{\ell=0}^p \Delta \tilde{x}_{i,t-\ell}(m) \beta_{\ell} + \varepsilon_{i,t} \\ \Delta r_{i,t} &= a_i - \underbrace{\left(1 - \sum_{\ell=1}^p \varphi_{\ell}\right)}_{\phi} r_{i,t-1} + \sum_{\ell=1}^{p-1} \gamma_{\ell} \Delta r_{i,t-\ell} + \sum_{\ell=0}^p \Delta \tilde{x}_{i,t-\ell}(m) \beta_{\ell} + \varepsilon_{i,t} \\ \Delta r_{i,t} &= a_i - \phi \left(r_{i,t-1} - \underbrace{\Delta \tilde{x}_{i,t}(m)}_{\theta} \sum_{\ell=0}^p \beta_{\ell} \right) + \sum_{\ell=1}^{p-1} \gamma_{\ell} \Delta r_{i,t-\ell} \\ &\quad - \sum_{\ell=1}^p \left(\Delta \tilde{x}_{i,t}(m) - \Delta \tilde{x}_{i,t-\ell}(m) \right) \beta_{\ell} + \varepsilon_{i,t} \\ \Delta r_{i,t} &= a_i - \phi \left(r_{i,t-1} - \Delta \tilde{x}_{i,t}(m) \theta \right) + \sum_{\ell=1}^{p-1} \gamma_{\ell} \Delta r_{i,t-\ell} + \sum_{\ell=0}^{p-1} \Delta^2 \tilde{x}_{i,t-\ell}(m) \delta_{\ell} + \varepsilon_{i,t}, \end{aligned} \tag{3}$$

where

- $\gamma_{\ell} = -(\varphi_{\ell+1} + \dots + \varphi_p)$, for $\ell = 1, \dots, p-1$;
- $\delta_{\ell} = -(\beta_{\ell+1} + \dots + \beta_p)$,

and Δ^2 denotes the second order difference of \tilde{x} . In this formulation of the model, the coefficients of interest are ϕ and θ , where θ captures the long-run effect of the climate variables on the growth rate, and ϕ captures the speed of adjustment to this long-run relationship. It is also worth noticing that the fixed effect a_i is still included in the regression, and that therefore this model cannot be directly estimated by Ordinary Least Squares (OLS).

Replication of Empirical Results

We first show that using the model in equation (3), i.e. the error correction version of the model, is numerically equivalent to estimate directly the model in equation (2), i.e., what I refer to as the *reduced-form* model.

Results are reported in Tables 1 and 2 for the half-panel jackknife fixed effect estimator (HPJ-FE) of Chudik, Pesaran, and Yang (2018). Notice that the values of the coefficients are numerically equivalent, but there is some discrepancy in the standard errors. This may due to the fact that Stata and R use different methods to compute standard errors.² However, the values obtained are not too far apart and do not change the qualitative conclusions of the paper.

```
#### 3 DATABASES VARY ACCORDING TO THE
#### THE VALUE OF m USED (20,30,40)
## MA 20
data("data_kahn.ma20",package="KMNPY.2021")
pdata.kahn.ma20 <- pdata.frame(data.kahn.ma20, index=c("isocode", "year"))
## MA 30
data("data_kahn.ma30",package="KMNPY.2021")
pdata.kahn.ma30 <- pdata.frame(data.kahn.ma30, index=c("isocode", "year"))
## MA 40
data("data_kahn.ma40",package="KMNPY.2021")
pdata.kahn.ma40 <- pdata.frame(data.kahn.ma40, index=c("isocode", "year"))
```

```
formula.ardl.tab11 <- as.formula(dgrowth ~ l1growth + dtemp_plus + dtemp_minus + dprecip_plus + l2c
```

```
hpj.p.tab11.ma20 <- hpj.fe(formula.ardl.tab11,pdata.kahn.ma20,"individual")
hpj.p.tab11.ma30 <- hpj.fe(formula.ardl.tab11,pdata.kahn.ma30,"individual")
hpj.p.tab11.ma40 <- hpj.fe(formula.ardl.tab11,pdata.kahn.ma40,"individual")
```

```
tab1spec1.ma20 <- ec.coef.se.dm(hpj.p.tab11.ma20)
tab1spec1.ma30 <- ec.coef.se.dm(hpj.p.tab11.ma30)
tab1spec1.ma40 <- ec.coef.se.dm(hpj.p.tab11.ma40)
```

```
formula.ardl.tab12 <- as.formula(dgrowth ~ l1growth + dtemp_plus + dtemp_minus + l1dgrowth + l2c
```

```
hpj.p.tab12.ma20 <- hpj.fe(formula.ardl.tab12,pdata.kahn.ma20,"individual")
hpj.p.tab12.ma30 <- hpj.fe(formula.ardl.tab12,pdata.kahn.ma30,"individual")
hpj.p.tab12.ma40 <- hpj.fe(formula.ardl.tab12,pdata.kahn.ma40,"individual")
```

```
tab1spec2.ma20 <- ec.coef.se.dm(hpj.p.tab12.ma20,numreg = 2)
tab1spec2.ma30 <- ec.coef.se.dm(hpj.p.tab12.ma30,numreg = 2)
tab1spec2.ma40 <- ec.coef.se.dm(hpj.p.tab12.ma40,numreg = 2)
```

Table 1: HPJ-FE Estimator with $m = \{20, 30, 40\}$ in Table 1 of Kahn et al. (2021) (error correction form, standard errors in brackets).

	Specification 1			Specification 2		
	MA 20	MA 30	MA 40	MA 20	MA 30	MA 40
$\hat{\theta}_{\Delta T_{i,t}^*(m)}$	-0.566 (0.21)	-0.894(0.292)	-1.072(0.373)	-0.572 (0.210)	-0.908 (0.291)	-1.105 (0.372)
$\hat{\theta}_{\Delta T_{i,t}^*(m)}$	-0.5 (0.252)	-0.783(0.386)	-0.909(0.492)	-0.508 (0.252)	-0.806 (0.386)	-0.954 (0.492)
$\hat{\theta}_{\Delta P_{i,t}^*(m)}$	-0.031 (0.361)	0.122(0.571)	-0.005(0.791)	~ (-)	~ (-)	~ (-)
$\hat{\theta}_{\Delta P_{i,t}^*(m)}$	-0.175 (0.429)	-0.320(0.681)	-0.595(0.856)	~ (-)	~ (-)	~ (-)
$\hat{\phi}$	0.603 (0.046)	0.603(0.046)	0.602(0.046)	0.604 (0.046)	0.604 (0.046)	0.604 (0.046)

```
## ESTIMATE ONLY THE HALF-PANEL JACKKNIFE BUT NOT THE FIXED EFFECT
pdata.kahn.ma20 <- createIagvars(pdata.kahn.ma20, ma = 20)
pdata.kahn.ma30 <- createIagvars(pdata.kahn.ma30, ma = 30)
pdata.kahn.ma40 <- createIagvars(pdata.kahn.ma40, ma = 40)
```

```
## LOAD DATA
pdata.kahn.tab5.ma20 <- subset(pdata.kahn.ma20,as.numeric(pdata.kahn.ma20$year) >= 61)
pdata.kahn.tab5.ma30 <- subset(pdata.kahn.ma30,as.numeric(pdata.kahn.ma30$year) >= 61)
pdata.kahn.tab5.ma40 <- subset(pdata.kahn.ma40,as.numeric(pdata.kahn.ma40$year) >= 61)
```

```
# DIRECT SPECIFICATION 1
formula.ardl.tab11dir <- as.formula(growth ~ l1growth + l2growth + l3growth + l4growth + dtemp.pl
```

```
hpj.p.tab11dir.ma20 <- hpj.fe(formula.ardl.tab11dir,pdata.kahn.ma20,"individual")
hpj.p.tab11dir.ma30 <- hpj.fe(formula.ardl.tab11dir,pdata.kahn.ma30,"individual")
hpj.p.tab11dir.ma40 <- hpj.fe(formula.ardl.tab11dir,pdata.kahn.ma40,"individual")
```

```
tab2spec1.ma20 <- sr.coef.se.dm(hpj.p.tab11dir.ma20)
tab2spec1.ma30 <- sr.coef.se.dm(hpj.p.tab11dir.ma30)
tab2spec1.ma40 <- sr.coef.se.dm(hpj.p.tab11dir.ma40)
```

```
formula.ardl.tab12dir <- as.formula(growth ~ l1growth + l2growth + l3growth + l4growth + dtemp.pl
```

```
hpj.p.tab12dir.ma20 <- hpj.fe(formula.ardl.tab12dir,pdata.kahn.ma20,"individual")
hpj.p.tab12dir.ma30 <- hpj.fe(formula.ardl.tab12dir,pdata.kahn.ma30,"individual")
hpj.p.tab12dir.ma40 <- hpj.fe(formula.ardl.tab12dir,pdata.kahn.ma40,"individual")
```

```
tab2spec2.ma20 <- sr.coef.se.dm(hpj.p.tab12dir.ma20,numreg = 2)
tab2spec2.ma30 <- sr.coef.se.dm(hpj.p.tab12dir.ma30,numreg = 2)
tab2spec2.ma40 <- sr.coef.se.dm(hpj.p.tab12dir.ma40,numreg = 2)
```

Table 2: HPJ-FE Estimator with $m = \{20, 30, 40\}$ in Table 1 of Kahn et al. (2021) (reduced form, standard errors in brackets).

	Specification 1			Specification 2		
	MA 20	MA 30	MA 40	MA 20	MA 30	MA 40
$\hat{\theta}_{\Delta T_{i,t}^*(m)}$	-0.566 (0.204)	-0.894(0.281)	-0.894(0.361)	-0.592 (0.205)	-0.788 (0.285)	-0.941 (0.365)
$\hat{\theta}_{\Delta T_{i,t}^*(m)}$	-0.5 (0.246)	-0.783(0.374)	-0.783(0.478)	-0.489 (0.232)	-0.826 (0.348)	-1.117 (0.442)
$\hat{\theta}_{\Delta P_{i,t}^*(m)}$	-0.031 (0.358)	0.122(0.56)	0.122(0.771)	~ (-)	~ (-)	~ (-)
$\hat{\theta}_{\Delta P_{i,t}^*(m)}$	-0.175 (0.431)	-0.320(0.659)	-0.320(0.857)	~ (-)	~ (-)	~ (-)
$\hat{\phi}$	0.603 (0.046)	0.603(0.046)	0.603(0.046)	0.604 (0.046)	0.604 (0.046)	0.604 (0.046)

Replication of Counterfactual Results

This section reproduces the results of Section 3 in Kahn et al. (2021). I do so by estimating a ARDL model as in equation (2), where the absolute difference between the temperature and its historical mean is the only exogenous regressor. The historical mean is estimated by using windows of size $m = \{20, 30, 40\}$.

Coefficients from this estimation for $m = 30$ are reported in Table 3, which should be compared to Table 5 in Kahn et al. (2021).

```
## ADD SCENARIOS FROM KAHN ET AL. (2021)
data("data_kahn.scen",package = "KMNPY.2021")
data("datatapp",package = "KMNPY.2021")
```

```
data.kahn.scen <- merge(datatapp,data.kahn.scen,by = "iso")
```

```
pdata.kahn.tab5.ma20 <- pdata.frame(merge(pdata.kahn.tab5.ma20,data.kahn.scen,by = "iso"),
index=c("isocode", "year"))
```

```
pdata.kahn.tab5.ma30 <- pdata.frame(merge(pdata.kahn.tab5.ma30,data.kahn.scen,by = "iso"),
index=c("isocode", "year"))
```

```
pdata.kahn.tab5.ma40 <- pdata.frame(merge(pdata.kahn.tab5.ma40,data.kahn.scen,by = "iso"),
index=c("isocode", "year"))
```

```
formula.ardl <- as.formula(growth ~ l1growth + l2growth + l3growth + l4growth +
dtempabs + l2dtempabs + l2dtempabs + l3dtempabs + l4dtempabs)
```

```
hpj.p.1.ma20 <- hpj.fe(formula.ardl,pdata.kahn.tab5.ma20,"individual")
hpj.p.1.ma30 <- hpj.fe(formula.ardl,pdata.kahn.tab5.ma30,"individual")
hpj.p.1.ma40 <- hpj.fe(formula.ardl,pdata.kahn.tab5.ma40,"individual")
```

Table 3: Table 5 in Kahn et al. (2021).

	$\hat{\psi}_1$	$\hat{\psi}_2$	$\hat{\psi}_3$	$\hat{\psi}_4$	$\hat{\psi}_5$	$\hat{\psi}_6$	$\hat{\psi}_7$	$\hat{\psi}_8$	$\hat{\psi}_9$
Coefficient	0.2843	0.0785	0.0547	-0.0016	-0.0038	-0.0056	-0.0094	-0.0090	-0.0060
Std. Err.	0.0497	0.0270	0.0221	0.0329	0.0021	0.0029	0.0031	0.0026	0.0021

We further proceed to obtain the expected losses in GPD by country for a given scenario about the expected evolution of temperatures around their historical mean.

To do so, we first rewrite equation (2) in the following form

$$(L)r_{i,t} = a_i + \beta(L)\Delta x_{i,t}(m) + \varepsilon_{i,t}, \tag{4}$$

where $\Delta x_{i,t}(m) = |T_{i,t} - T_{i,t-m}^*(m)|$,

$$\varphi(L) = 1 - \sum_{\ell=1}^p \varphi_{\ell} L^{\ell}$$

$$\beta(L) = \sum_{\ell=0}^p \beta_{\ell} L^{\ell},$$

and L is the lag operator. That is,

$$r_{i,t} = \varphi(L)^{-1} a_i + \psi(L) \Delta x_{i,t}(m) + \theta(L) \varepsilon_{i,t},$$

where $\theta(L) = \sum_{\ell=0}^{\infty} \theta_{\ell} L^{\ell}$, and

$$\psi(L) = \varphi(L)^{-1} \beta(L) = \sum_{\ell=0}^{\infty} \psi_{\ell} L^{\ell}.$$

Obtaining an estimator of ψ allows us to forecast the expected value of GDP growth h periods ahead, for a given sequence of observed temperatures (in deviations from their historical mean) between T and $T+h$.

In Kahn et al. (2021), the authors consider that temperature follows a model with a linear trend plus noise. Under an additional assumption that the noise is normally distributed, they derive the expected change in temperature between time T and time $T+h$ in closed form, and use it to predict the expected loss in GDP for several scenarios. In particular, they have

$$T_{i,T+j} = a_{Ti} + b_{Ti,j}(T) + v_{Ti,T+j}, \text{ for } j = 1, 2, \dots$$

$$v_{Ti,j} \sim N(0, \sigma_{Ti}^2)$$

$$E[T_{i,T+j} - T_{i,T+j-1}^*(m)] = \mu_{Ti,j} \left[\Phi \left(\frac{\mu_{Ti,j}}{\omega_{Ti}} \right) - \Phi \left(-\frac{\mu_{Ti,j}}{\omega_{Ti}} \right) \right] + 2\omega_{Ti} \phi \left(\frac{\mu_{Ti,j}}{\omega_{Ti}} \right),$$

where $\mu_{Ti,j} = b_{Ti,j}(m+1)/2$, $\omega_{Ti}^2 = \sigma_{Ti}^2(1+1/m)$, and $\{\phi, \Phi\}$ are the pdf and cdf of a standard normal distribution.

Let $g_{Ti}(m, b_{Ti,j}) = E[T_{i,T+j} - T_{i,T+j-1}^*(m)]$, the change in the log real GDP per-capita of country i over the horizon h , is given by

$$\Delta_h(d_i, m) = \sum_{j=1}^h \psi_{h-j} \left(g_{Ti}(m, b_{Ti}^0 + j d_i, \sigma_{Ti}^0) - g_{Ti}(m, b_{Ti}^0, \sigma_{Ti}^0) \right),$$

where b_{Ti}^0 and σ_{Ti}^0 are the trend and the standard deviation of the error term estimated using observations from years 1960-2014.

Kahn et al. (2021) consider the uncertainty coming from the estimation of ψ , and provide confidence intervals for these parameters. However, they do not provide forecast intervals for the expected loss in GDP $\hat{\psi}$ is a nonlinear transformation of $\{\hat{\phi}, \hat{\beta}\}$, so that standard asymptotic confidence intervals based on asymptotic normality of the estimator are hard to derive in closed form. In order to obtain confidence intervals, I bootstrap the panel with respect to its cross sectional dimension using a dynamic wild bootstrap to allow for serial dependence between countries (Shao (2010), and Gao, Pen, and Yan (2022)). The first 20 values of $\hat{\psi}$ with their 95% confidence intervals are given in Figure 1, which mimics very closely Figure 5 in the original paper.

We present here an example with $m = 30$ about how to invoke the function to obtain 95% bootstrap confidence intervals with $B = 499$ replications.

```
## SET PARAMETERS FOR BOOTSTRAP
set.seed(123)
B <- 499
```

```
psihat.b.ma30 <- psicoef.b(pdata.kahn.tab5.ma30, formula.ardl,
hpj.p.1.ma30$coeff, hpj.p.1.ma30$resid, B = B)
```

```
psihat025.ma30 <- apply(psihat.b.ma30$psihat,1,function(x) quantile(x,0.025))
psihat975.ma30 <- apply(psihat.b.ma30$psihat,1,function(x) quantile(x,0.975))
```

```
barCenters <- barplot(psihat.ma30[1:21],xlim = c(0,21),ylim = rev(c(-0.02,0.001)),col = "blue4",c
```

```
segments(barCenters, psihat025.ma30[1:21], barCenters,
psihat975.ma30[1:21], lwd = 2,col = "red4")
```

```
arrows(barCenters, psihat025.ma30[1:21], barCenters,
psihat975.ma30[1:21], lwd = 2, angle = 90,
```

```
code = 3, length = 0.1,col = "red4")
```

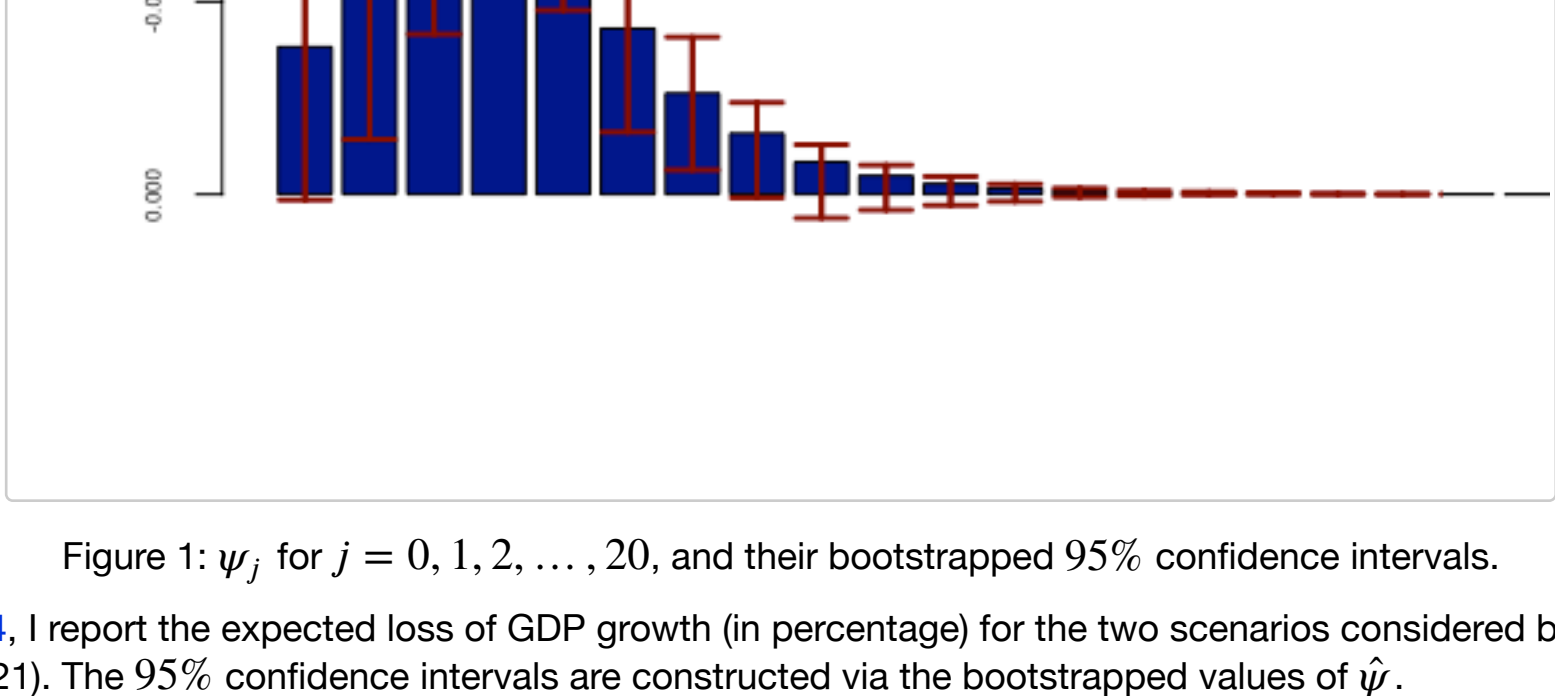


Figure 1: ψ_j for $j = 0, 1, 2, \dots, 20$, and their bootstrapped 95% confidence intervals.

In Table 4, I report the expected loss of GDP growth (in percentage) for the two scenarios considered by Kahn et al. (2021). The 95% confidence intervals are constructed via the bootstrapped values of $\hat{\psi}$.

```
m.vec <- seq(20,40,10)
H <- 85
# Use q specific sequence of values for the change in trend
d.vec <- seq(min(climate.data2014$d1_rcp26),max(climate.data2014$d1_rcp85),length.out = 50)
```

```
b.med <- mean(climate.data2014$beta.T1)
s.med <- mean(climate.data2014$sigma.T1)
```

```
i <- 1
delta.diff.med.1 <- matrix(0,3,50)
for(jj in 1:50){
# M = 20
delta.diff.med.1[1,jj] <- cccounter(m.vec[1],b.med[1],s.med[1],1,psihat.ma20,d.vec[jj])$delta[2]
# M = 30
delta.diff.med.1[2,jj] <- cccounter(m.vec[2],b.med[1],s.med[1],1,psihat.ma30,d.vec[jj])$delta[2]
# M = 40
delta.diff.med.1[3,jj] <- cccounter(m.vec[3],b.med[1],s.med[1],1,psihat.ma40,d.vec[jj])$delta[2]
}
```

```
delta.diff.med.1 <- t(delta.diff.med.1)
```

```
colnames(delta.diff.med.1) <- paste0("m = ",m.vec)
```

```
nn <- 3
layout(matrix(c(1,2),nrow=1), width=c(12,3))
```

```
par(mar=c(5,4,4,0)) #No margin on the right side
```

```
matPlot(d.vec,delta.diff.med.1, type = "l",xlab = bquote(d[i]),ylab = bquote(Delta[h] ~ "(C" ~ d[i
```

```
par(mar=c(5,0,4,2)) #No margin on the left side
```

```
plot(c(0,1),type="n", axes=F, xlab="", ylab="")
```

```
legend("center", colnames(delta.diff.med.1),col=seq_len(nn),fill=seq_len(nn))
```

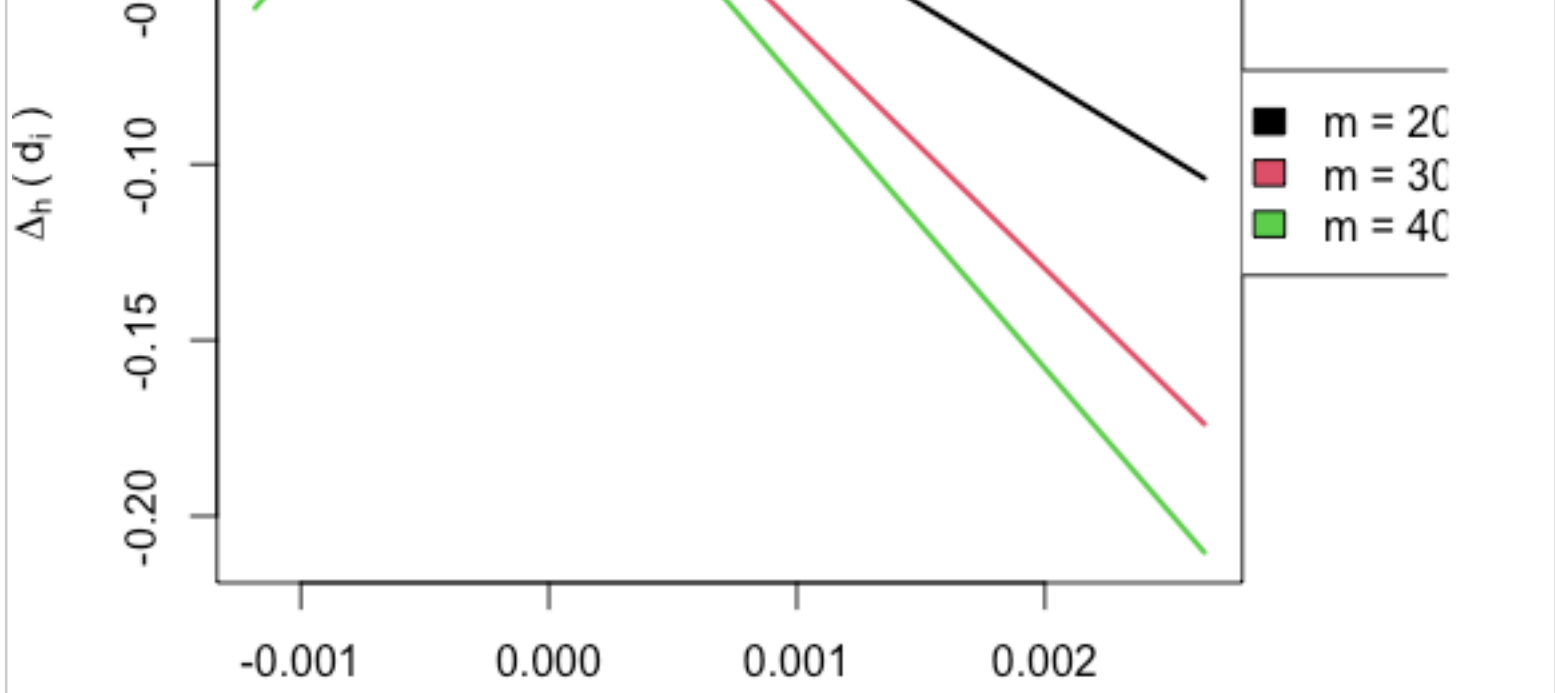


Figure 2: Counterfactual scenarios for several values of d_i and $m = \{20, 30, 40, 50\}$.

Figure 2 showcases the expected change in log GDP per capita as a function of the parameter d_i , and for $m = \{20, 30, 40\}$.

Table 4: Table 5: Table 6 in Kahn et al. (2021) with $m = 30$.

~	2030	2050	2100
World			
RCP 2.6	-0.01	0.11	1.07
RCP 8.5	0.80	2.51	7.22
China			
RCP 2.6	-0.45	-0.80	0.45
	[-0.21,0.01]	[-0.37,0.01]	[-0.00,0.88]
RCP 8.5	0.58	1.62	4.35
	[0.27,0.88]	[0.74,2.38]	[2.00,6.38]
European Union			
RCP 2.6	-0.05	-0.04	0.45
	[-0.13,0.03]	[-0.22,0.15]	[0.06,0.82]
RCP 8.5	0.53	1.70	5.25
	[0.24,0.80]	[0.78,2.51]	[2.42,7.71]
India			
RCP 2.6	0.26	0.61	2.57
	[0.12,0.38]	[0.37,1.20]	[1.16,3.77]
RCP 8.5	1.16	3.62	9.90
	[0.53,1.78]	[1.67,5.34]	[4.55,14.52]
Russia			
RCP 2.6	-0.14	-0.34	-0.71
	[-0.21,-0.08]	[-0.50,-0.18]	[-1.03,-0.32]
RCP 8.5	1.03	3.08	8.93
	[0.47,1.55]	[1.42,4.54]	[4.11,13.10]
United States			
RCP 2.6	0.20	0.60	1.88
	[0.09,0.30]	[0.28,0.86]	[0.87,2.76]
RCP 8.5	1.20	3.77	10.52
	[0.55,1.80]	[1.74,5.56]	[4.84,15.42]

References

Chudik, Alexander, M. Hashem Pesaran, and Jui-Chung Yang. 2018. "Half-Panel Jackknife Fixed-Effects Estimation of Linear Panels with Weakly Exogenous Regressors." *Journal of Applied Econometrics* 33 (6): 816–36. <https://doi.org/10.1002/jae.2623>

Gao, Jiti, Bing Pen, and Yayi Yan. 2022. "A Simple Bootstrap Method for Panel Data Inferences." *Monash University Department of Econometrics and Business Statistics WP Series*.