




Intratemporal elasticity of substitution between private and public consumption: new evidence and implications

John Nana Francois¹ · Andrew Keinsley² 

Received: 16 June 2022 / Accepted: 9 February 2023 / Published online: 7 April 2023

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract

This paper estimates the intratemporal elasticity of substitution (IES) between private and public consumption in private utility. Using panel data for 17 European countries over 1970–2018, we estimate the IES to lie between 0.6 and 0.74. When combined with the relevant intertemporal elasticity of substitution, our estimated IES implies that private and public consumption are Edgeworth complements. The panel estimate, however, conceals a large heterogeneity where the IES ranges from as low as 0.3 in Italy to 1.3 in Ireland. This suggests that the crowding-in (out) effect of fiscal policies involving change in government consumption will differ across countries. This cross-country variation in the IES is positively correlated the share of health spending in public expenditure but negatively correlated with the share of public order and safety in government expenditures. We also find a U-shape relationship between the size of IES and government size.

Keywords Intratemporal elasticity of substitution · Heterogeneity · Edgeworth complementarity · Public and private consumption · Fiscal policy

JEL Classification C22 · C23 · E62 · H5

We have NO conflicts of interest to disclose. We would like to thank the editor and the two anonymous referees for important comments and suggestions that have significantly improved the paper. We are also grateful to Stephen A. O’Connell, Maya Eden, and Anson Zhou for useful feedback, discussions, and comments. Finally, we thank conference participants at the 90th Southern Economic Association Annual Meeting in New Orleans and the 2020 conference series at the International Research Department at the Reserve Bank of India for useful comments and feedback. All errors are hereby ours. **Disclaimer:** The findings, interpretations and conclusions expressed in this paper are entirely those of the authors and should not be attributed to the World Bank, its Executive Directors, or the countries they represent.

✉ Andrew Keinsley
andrewkeinsley@weber.edu

John Nana Francois
jfrancois1@worldbank.org

¹ Fiscal Policy and Sustainable Growth Unit, World Bank, Washington, USA

² Weber State University, Ogden, USA

1 Introduction

The idea that a macroeconomic model of consumption should allow for the direct effect of government consumption in private utility is standard (Kormendi 1983; Evans and Karras 1996; Amano and Wirjanto 1997, 1998; Fève et al. 2013; Bouakez and Rebei 2007; Leeper et al. 2017; Ganelli and Tervala 2009; Fiorito and Kollintzas 2004). Here, whether private and public consumption are Edgeworth complements or substitutes becomes critical in mediating fiscal policies involving changes in government consumption. Specifically, through a positive marginal utility channel, an increase in government consumption can directly increase private consumption if the two consumption goods are Edgeworth complements. This is particularly important because if the degree of complementarity is high enough, the positive marginal utility of private consumption can offset, and possibly outweigh the standard negative wealth effect arising from financing the increase in public consumption with taxes. In contrast, Edgeworth substitutability between private and public consumption generates the opposite effect, suggesting that cuts in public consumption can induce a demand-side offset that can further lead to some moderation in the impact of fiscal consolidation on output.

This paper provides a fresh look into the empirical relationship between private and public consumption in household utility by estimating the intratemporal elasticity of substitution for 17 European countries over the period 1970–2018. In the context of the CES utility function, and for a given intertemporal elasticity of substitution, the intratemporal elasticity of substitution (IES) between private and government consumption plays a central role in determining whether the two goods are Edgeworth complements or substitutes.¹ However, existing studies that estimate the relationship between private and government consumption in the panel data framework have done so under the “strong” assumption of cross-sectional independence across countries (e.g., Fiorito and Kollintzas 2004; Kwan 2009; Ho 2001; Dawood and Francois 2018; Brown and Wells 2008; Jalles and Karras 2021). Specifically, these studies have neglected the presence of cross-sectional dependencies that may exist across countries either due to global shocks (e.g., oil price shocks and the ongoing COVID-19 pandemic, which calls for almost synchronized policy actions across the world) or economic spillovers (e.g., provision of public goods in neighboring countries, economic integration or fiscal policy coordination between countries (e.g., Banerjee and Carrion-i Silvestre 2017)). Neglecting these dependencies can lead to a breakdown of crucial assumptions for standard panel estimators employed in existing cross-country studies, which can induce biased estimates and spurious inferences (Chudik et al. 2017; Eberhardt and Teal 2020).

Additionally, the cointegration techniques employed in existing studies for both time series and panel data provide very little guidance on model specification in uncovering the IES. More precisely, with a common underlying theory to motivate the estimation, some studies have used the price of the consumption goods as the regressand (e.g., Amano and Wirjanto 1998) while other studies instead employ consumption as the regressor (e.g., Kwan 2009; Ho 2001; Dawood and Francois 2018). The choice of regressor and regressand is mixed across studies partly because the intratemporal

¹ I illustrate the relevance of intratemporal elasticity of substitution in a simple RBC model in section 2.

equilibrium condition from the model equates the consumption ratio to their inverse price ratio. The researcher can therefore choose any of the two variables as the dependent variables to recover the IES, and in theory, one should expect to recover the same IES regardless of the choice of dependent variable. Unfortunately, existing literature does not provide any evidence to show whether the choice of one dependent variable over the other can impact the size of the IES due model misspecification and hence, lead to incorrect inference. Indeed, Ng and Perron (1997) show that the choice of variable to put on the left-hand side as the regressand affects the precision of estimates and can lead to drastically different point estimates; hence, a careful selection of the regressand is required. Importantly, the prior literature that employs panel data and cointegration techniques fails to provide detailed/useful post-estimation diagnostics of well-behaved residuals to ensure that models are not misspecified and hence, the estimates of the IES valid.

In this paper, we explore the use of panel cointegration techniques and recently developed Cross-section augmented distributed lag (CS-DL) estimators in uncovering the IES while accounting for the aforementioned shortcomings in the existing literature. Specifically, for the empirical work, we remain agnostic by not selecting a fixed set of assumptions *a priori*. Consequently, in the baseline analysis we estimate the IES under: (1) different choices of estimators, (2) different model specifications in that the regressor and regressand in the cointegration relation are either chosen as the price or consumption ratio, and (3) cross-sectional dependence *versus* independence. Furthermore, as a means of dealing with cross-sectional dependence, we employ two strategies—one that implements the time-demeaned variable approach *à la* Sarafidis et al. (2009), Herzer and Morrissey (2013) and Francois and Keinsley (2019), and an alternative that augments cross-sectional averages of all the variables to the estimation as in Pesaran (2006), Holly et al. (2010), and Eberhardt and Teal (2020). Avoiding a preset of assumptions allows us highlight how models in existing studies might have been misspecified. We therefore provide systematic post-estimation diagnostics in the manner of Eberhardt and Presbitero (2015) and Chudik et al. (2017) to test these sets of assumptions, evaluate model specifications, validate the estimated IES. These features are not shared with other studies that estimate the IES (Kwan 2009; Ho 2001; Dawood and Francois 2018; Brown and Wells 2008; Amano and Wirjanto 1997).²

The econometric methodology considered in this paper is based on a cointegrating regression model interpreted as an equilibrium condition in an optimal fiscal policy setting. The CES aggregator function is used to define effective consumption, which is a function of private and public consumption. The following results emerge from the baseline estimations: First, all estimates of the IES support a direct role of public consumption in private utility, suggesting that public consumption is utility enhancing in European countries. These results are consistent with findings in the literature (see,

² Unlike Fiorito and Kollintzas (2004) where no temporal utility function is specified, we specify a CES aggregator function for effective consumption. The study, however, shares the same underlying assumption of permanent income hypothesis. Additionally, while Fiorito and Kollintzas (2004) focuses on the sub-categories of government consumption (i.e., merit and public goods), we concentrate on aggregate public consumption, which is often focused on in both empirical and theoretical work. We, however, employ the decomposition of public consumption into merit and public goods to attempt to explain the relationship between private and government consumption studied in the paper.

Leeper et al. 2017; Fiorito and Kollintzas 2004, for instance). Second, depending on the set of assumptions made prior to estimation, the estimated IES can range from as low as 0.3 to as high 3.3, implying gross complementarity and substitutability between private and public consumption for the same dataset. Specifically, in the estimation where the consumption ratio between private and public consumption is used as the regressand, we find an estimated IES that is between 0.297—0.901. These estimates consistently uncover that government and private consumption are *gross complements*. In contrast, in the case where the price ratio of the consumption goods is employed as the regressand, we find estimates of the IES that are consistently greater than 1 implying *gross substitutability* between private and public consumption. Fourth, post-estimation diagnostics in the error structure of the regressions, which is often ignored in this literature, uncover that estimates from the consumption ratio regressions outperform the case where the price ratio is employed as the regressand, and that the most reliable estimates of the IES lie between 0.6 and 0.74. Importantly, this value, when combined with the relevant intertemporal elasticity of substitution, implies *Edgeworth complementarity* between private and public consumption—a finding that supports studies such as Fiorito and Kollintzas (2004) and Jalles and Karras (2021). Interestingly, these set of values of the IES are able to predict a positive response of private consumption in standard RBC, which coincides with results from empirical models (e.g., Leeper et al. 2017; Laumer 2020; Ilzetzki et al. 2013). Finally, we recognize that point estimates from the panel analysis can mask potential cross-country heterogeneity in the IES. Consequently, we carry out a country-by-country analysis. The results from the heterogeneous panel estimates reveal that the IES can range from as high as 1.2 in Ireland to as low as 0.3 in Italy. The variation observed in the size of the estimate is positively correlated with the share of health spending in total government expenditure but negatively associated with the share of public order expenditure in total government spending. We also find a U-shape relationship between the size of IES and government size.

The focus on European economies is particularly important and timely in that with weak economic growth in Europe, the idea is being resurrected that European governments should pursue fiscal stimulus by increasing spending—i.e., public consumption and investment spending.³ The findings in this paper shed light into the effectiveness of fiscal policy in these economies by considering the marginal utility channel of public consumption. That is, the results that public and private consumption are Edgeworth complements in European countries suggest that expansionary fiscal policy that targets increasing government consumption can induce Keynesian effects. This is because with Edgeworth complementarity, an increase in government consumption raises the marginal utility of private consumption. If this increase in marginal utility is strong enough, it can outweigh the traditional wealth effect that arises from financing the increase in public consumption. Furthermore, given that current monetary policy in European economies is accommodative of fiscal policy at least in the foreseeable future, these Keynesian effects produced by the marginal utility channel of public consumption are likely to be amplified. Additionally, the findings in this paper

³ See for instance, <https://www.aicgs.org/2020/01/fiscal-stimulus-for-the-european-economy-the-right-thing/> for further discussions.

suggest that fiscal stimulus packages that comprise large government consumption components may be effective at stimulating aggregate demand, reinforcing similar conclusions in the literature (see for instance, Boehm 2019).⁴ In a symmetric thinking, the results suggest that with public and private consumption being Edgeworth complements, fiscal consolidation can be self-defeating. This finding is isomorphic to conclusions in Bandeira et al. (2018), who show that fiscal consolidation in the form of cuts in productive public good can be contractionary.

The rest of the paper is organized as follows: Sect. 2 presents some stylized facts on to the empirical- and model-based response of private consumption following an increase in government consumption with emphasis on the relevance of the IES. Section 3 motivates the empirical analysis with a simple theoretical model interpreted as an equilibrium condition in an optimal fiscal policy setting. Section 4 presents the econometric methodology. Section 5 discusses the baseline results and their implications. Section 6 explores the possibility of heterogeneity in the estimated values. Section 7 concludes.

2 Some stylized facts

In this section, we perform two simple but important exercises to highlight some stylized facts. First, we present the empirical response of private consumption to an increase in public consumption using panel vector autoregression (VAR) model. We apply the panel data of the 17 European countries over the period 1970–2017. Second, we illustrate the relevance of the size of the IES in replicating (at least in a qualitative manner) the empirical response of private consumption to an increase in public consumption. We employ a calibrated real business cycle (RBC) model for this exercise. In both cases, we present the results via an impulse response of private consumption to an exogenous increase to public consumption.⁵

2.1 What Does The Data Say?

Here, we provide a simple exercise to highlight the response of private consumption to a government consumption increase. Uncovering this empirical response would help us appreciate the implications of the estimated intratemporal elasticities in Sect. 4, which is the main aim of the paper. To this end, we consider a reduced-form bivariate panel VAR of order p with panel-specific fixed effects represented by the following system of linear equations,

$$\mathbf{Y}_{it} = \mathbf{u}_i + \sum_{j=1}^p \mathbf{Y}_{it-j} \mathbf{A}_j + \mathbf{e}_{it} \quad (1)$$

⁴ Boehm (2019) shows that the government investment multiplier is near zero whereas the government consumption multiplier is approximately 0.8.

⁵ The section also allows for a partial review of the literature.

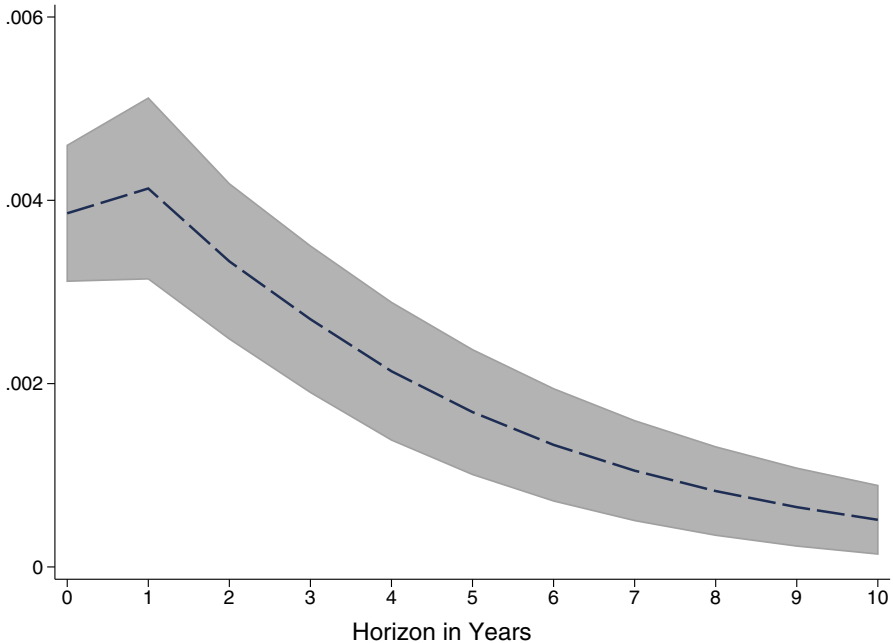


Fig. 1 Response of Private Consumption to an increase in Government Consumption. **Notes:** The shaded grey area is the 68% confidence interval. The lag order p is set to two according to the modified Akaike information criteria and Bayesian information criteria. The first four lags of the endogenous variables in the system are used as instruments

\mathbf{Y}_{it} is a $(1 \times k)$ vector of variables with $k = 2$ and it comprises the stationary versions of data on government consumption and private consumption for a given year t and country i . \mathbf{A}_j 's are $(k \times k)$ matrices to be estimated and \mathbf{u}_i and \mathbf{e}_{it} are $(1 \times k)$ vectors of dependent variable-specific panel fixed-effects and idiosyncratic errors, respectively. In the manner of Holtz-Eakin et al. (1988) and Abrigo and Love (2016), we assume that individual countries in the panel share the same underlying data generating process, with the reduced form parameters \mathbf{A}_j 's common across countries. Systematic cross-sectional heterogeneity is modeled as panel-specific fixed effects. The estimation of the parameters of the system in (1) is conducted in a generalized method of moments framework (see, Abrigo and Love 2016, for details).⁶ To recover the structural shocks from the VAR innovation in this simple bivariate setting, we adopt the recursive Cholesky decomposition identification, where government consumption is ordered first.

The response of private consumption to an increase in government consumption via impulse response function is depicted in Fig. 1. It is evident from Fig. 1 that an exogenous increase in government consumption increases private consumption, and

⁶ I employ the Stata program developed by Abrigo and Love (2016) for this simple exercise. As stressed by the authors, the GMM estimator suffers from weak instrument problems when the endogenous variables have a unit root. we run unit-root tests on the differenced data and find no unit roots in the variables we employ for the exercise.

this positive response is statistically significant. Although the results in Fig. 1 provide a preliminary relationship, the positive response of private consumption is mostly in line with studies such as Ilzetzi et al. (2013), Jha et al. (2014), Kilponen et al. (2019), and Laumer (2020) who find similar response of private consumption following an increase in government consumption.

2.2 How Relevant Is The IES?

The findings in the previous section and evidence in the empirical literature reinforce the idea that a positive consumption multiplier is not uncommon.⁷ However, this increase in private consumption following an increase in government consumption is at odds with predictions from neoclassical macroeconomic theory. Specifically, in standard RBC models, an increase in public consumption induces a negative wealth effect from financing the rise in government consumption with taxes, which reduces the household's permanent income (Baxter and King 1993). This drop in income forces household's to increase their labor supply; however, this positive labor supply is not strong enough to offset the negative wealth effect, and consumption falls in equilibrium. Additionally, although standard New Keynesian models improve the negative wealth effect, they do not generate positive consumption multipliers.⁸ To reconcile model-based predictions to empirical consumption multipliers, the New Keynesian framework, which features sticky prices, is extended to include non-Ricardian households (Galí et al. 2007).

The introduction of the non-Ricardian households breaks the standard Ricardian equivalence present in the baseline RBC model since this fraction of household agents cannot borrow or save and hence, consume all their current income. Consequently, because wages rise after an increase in government consumption, these non-optimizing households raise their consumption. If the fraction of non-Ricardian households is large enough, this can generate a net positive response for consumption following the increase in government consumption (See, Leeper et al. 2017).⁹

An additional, and perhaps, alternative and intuitive transmission mechanism, which does not require a rule-of-thumb household assumption, that has been proposed in the model-based literature is the complementarity between private and public

⁷ There is, however, a number of studies that do not find a positive response of private consumption following an increase in government consumption. See for instance, Ramey and Shapiro (1998) and Ramey (2011). Other studies instead find mixed response of private consumption following an increase in government consumption (see Hur et al. 2010, for the case of developing Asia)

⁸ In standard New Keynesian models with zero lower bound of interest rates, a rise in public consumption induces fiscal multipliers greater than unity (Christiano et al. 2011)

⁹ In a far from an exhaustive list, other existing studies such as Coenen and Straub (2005), Forni et al. (2009), and Hagedorn et al. (2019) extend Galí et al. (2007) to include other mechanisms to generate a positive consumption multiplier. The fraction of non-Ricardian household in these studies has been set between 0.2–0.5. Recently, however, Havranek and Sokolova (2020) find that the fraction of non-Ricardian household in the data is small (i.e., 0.11) and therefore, implies little evidence of deviations from consumption smoothing. Other studies such as Davig and Leeper (2011) and Zubairy (2014) explore other channels (i.e., regime-dependent multipliers and deep habits, respectively) away from the one proposed by Galí et al. (2007) to generate a positive multiplier. Other studies such as Canova and Pappa (2011) investigate the theoretical conditions for effectiveness of government consumption expenditure expansions.

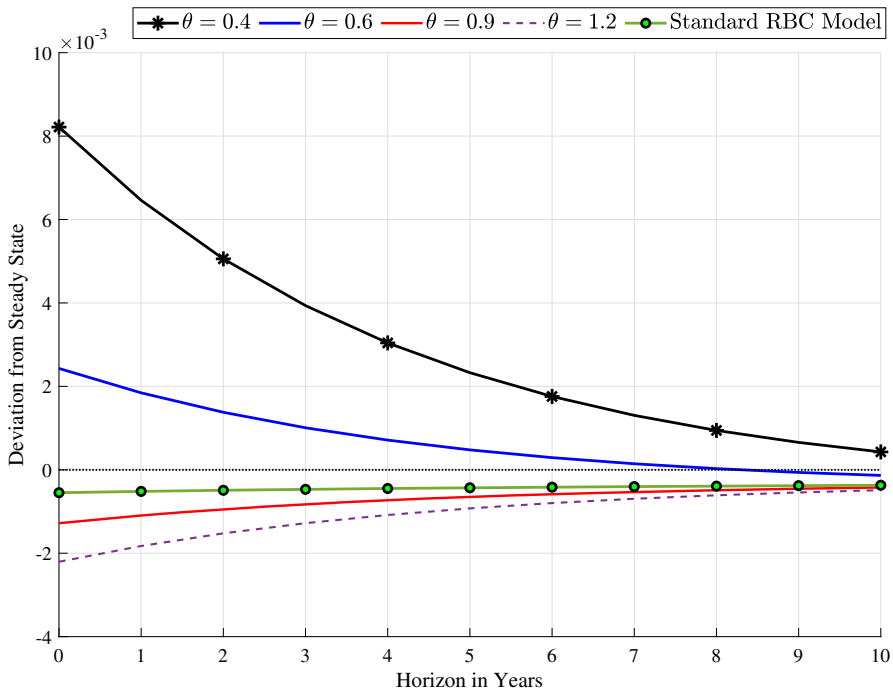


Fig. 2 Model-Based Response of Private Consumption to an increase in Government Consumption. **Notes:** Model-based response of private consumption following an increase in government consumption in a standard RBC model with utility-enhancing public consumption and capital adjustment costs. Consumer utility is the standard CRRA with $U = (C^e)^{1-1/\gamma} / (1-1/\gamma)$, where the intertemporal elasticity is given as $1/\gamma = 0.8$ and effective consumption (C^e) is defined as $C_t^e = [\lambda \varepsilon_t C_t^{1-(1/\theta)} + (1-\lambda) \varepsilon_t G_t^{1-(1/\theta)}]^{1/(1-(1/\theta))}$. The parameter θ governs the IES. $sgn[U_{C,G}] = [1/\gamma - \theta]$ governs the Edgeworth complementarity (substitutability) between private and government consumption where $U_{C,G} > 0 (< 0)$ represents complementarity (substitutability). The case where $U_{G,C} = 0$ is the standard neoclassical case where only wealth effect exists. With standard calibration of other model parameters, we still find a positive response of private consumption (on impact) to an increase in government consumption for a value of θ up to approximately 0.74

consumption.¹⁰ Here, when government consumption is a complement to private consumption in an Edgeworth–Pareto sense, an increase in government consumption raises the marginal utility of private consumption and provides additional motives for households to work more. When the degree of complementarity is sufficiently high, this positive marginal utility channel can offset the standard negative wealth effect and induce a positive consumption response. In a model that assumes a CES aggregator form for effective consumption, the intratemporal elasticity of substitution becomes a central parameter in driving the *size* and even the *sign* of the consumption response.

¹⁰ It is important to mention that the presence of this channel does not discount the other channels discussed in the literature such as the presence in incomplete markets (Hagedorn et al. 2019), heterogeneous agents (e.g., Achdou et al. 2017; Galí et al. 2007; Hagedorn et al. 2019), deep habits (e.g., Zubairy 2014), and regime e-switching policy (e.g., Davig and Leeper 2011). The marginal utility channel complements these other channels in expanding our understanding on the size and sign of the government consumption multiplier.

To elucidate the relevance of the intratemporal elasticity of substitution, we embed utility-enhancing government consumption in a basic RBC model. Notice that RBC models are the bedrock of most general equilibrium models and in its pure form, it produces a negative consumption effect in response to an increase in government consumption. Hence, using an RBC model provides a good starting point to highlight the relevance of this simple modification. To this end, the utility function is specified as a standard CRRA utility function with the intertemporal elasticity of substitution defined as $1/\gamma$ and effective consumption, C^e defined as $C^e = [\lambda C_t^{1-(1/\theta)} + (1 - \lambda)G_t^{1-(1/\theta)}]^{1/(1-(1/\theta))}$, where θ is the intratemporal elasticity of substitution, λ is the weight of private consumption in effective consumption and C_t and G_t are private and government consumption, respectively.¹¹ We fix the intertemporal elasticity of substitution associated with the utility function to be 0.8 as in Havránek (2015) for this exercise. Havránek (2015) discusses that 33 studies published in the top five general interest journals report an intertemporal elasticity of substitution to be 0.9 on average. However, we employ the more conservative value of 0.8 as Havránek (2015) proceeds to show that calibration above 0.8 is inconsistent with the empirical literature. We present the response of private consumption following an exogenous increase in government consumption in Fig. 2 via impulse response functions.

As can be seen from Fig. 2, for the given intertemporal elasticity of substitution, the response of private consumption is inversely related to the size of the IES. That is, smaller values of the IES that ensure that private and government consumption are *Edgeworth complements* improves the negative response of private consumption in the standard RBC model. If the degree of complementarity is high enough (i.e., low θ), it can completely offset, and even outweigh the standard negative wealth effect inducing a positive response of private consumption. In contrast, when private and public consumption are *Edgeworth substitutes* the negative wealth effect is reinforced. This is evident in Fig. 2 as the impulse response function associated with larger IES generate a larger negative response of private consumption compared to the standard RBC model. The results from this exercise indicate that estimating the IES accurately from data with the cointegration techniques this paper aims to utilize is important as it can have direct impact on how private consumption responds to changes in public consumption.

3 A simple theory to guide the empirical model

In this section, we lend a structural interpretation to the empirical estimation that follow in Sect. 4. We borrow heavily from Amano and Wirjanto (1998), Kwan (2009), and Dawood and Francois (2018) and assume a representative agent (i.e., social planner) who gains utility from two goods, private and public. The agent's expected lifetime

¹¹ The model in this paper is similar to Bouakez and Rebei (2007); however, we leave the presence of habits out as we are not after the hump-shaped response of private consumption. Habits are explored in Bouakez and Rebei (2007).

utility function is governed by Eq. (2) and is subject to stationary preference shocks:

$$U_t = E_0 \sum_{t=0}^{\infty} \beta^t u(C_t^e, \epsilon_t, v_t), \tag{2}$$

where $u(\cdot)$ takes the constant relative risk aversion (CRRA) form $u(C^e) = \frac{(C^e)^{1-\frac{1}{\gamma}}}{1-\frac{1}{\gamma}}$, with $1/\gamma$ representing the intertemporal elasticity of substitution. Effective consumption C^e is a constant elasticity of substitution (CES) aggregate of private and public consumption:

$$C_t^e = [\lambda \epsilon_t C_t^{1-(1/\theta)} + (1 - \lambda) v_t G_t^{1-(1/\theta)}]^{1/(1-(1/\theta))}, \tag{3}$$

where the random preference shocks (ϵ_t, v_t) are strictly stationary with unit means. These stationarity assumptions imply that preferences are stable in the long run. The preference parameters $\lambda \in [0, 1]$ and $\theta > 0$ represent the relative weight assigned to private goods and the intratemporal elasticity of substitution (IES), respectively. The latter restriction ensures that the standard assumption of convexity of preferences is preserved, a negative value of θ therefore violates this preference assumption. An intratemporal elasticity of substitution that is greater (less) than one implies gross substitutability (complementarity) between private and public consumption. When θ is equal to zero, the two goods are perfect gross complements. Finally, estimated values of θ less than zero are theoretically implausible as they violate standard properties of the consumer utility function (Ogaki et al. 1996). The agent maximizes her utility subject to the budget constraint $P_t^g G_t + P_t^c C_t = I_t$ where I_t is income.

With the assumption that the agent’s utility function is time-separable, the optimal consumption bundle satisfies an equality condition between the marginal rate of substitution (MRS) and the relevant relative price.¹² This yields the intra-temporal Euler equation of private versus public consumption for the social planner. Hence, we obtain the condition:

$$\frac{\partial U_t / \partial G_t}{\partial U_t / \partial C_t} \equiv \frac{v_t (1 - \lambda) C_t^{1/\theta}}{\epsilon_t \lambda G_t^{1/\theta}} = \frac{P_t^g}{P_t^c}. \tag{4}$$

Taking logs in Eq. (4), we obtain:

$$\ln \left(\frac{C_t}{G_t} \right) = -\theta \ln \left(\frac{1 - \lambda}{\lambda} \right) + \theta \ln \left(\frac{P_t^g}{P_t^c} \right) - \theta \ln \left(\frac{v_t}{\epsilon_t} \right). \tag{5}$$

As mentioned earlier, stability of preferences implies that the residual term $-\theta \ln(v_t/\epsilon_t)$ is stationary, and thus that Eq.(5) is a cointegrating regression provided that the log price ratio $\ln(P_t^g/P_t^c)$ and the log consumption ratio $\ln(C_t/G_t)$ are both

¹² The first-order conditions for private and public consumption from the optimization problem are $\partial U_t / \partial C_t = \epsilon_t \lambda C_t^{-1/\theta} C_e^{1/\theta} u'$ and $\partial U_t / \partial G_t = v_t (1 - \lambda) G_t^{-1/\theta} C_e^{1/\theta} u'$, respectively.

I(1) processes. The combination of stable preferences and the optimality condition in Eq. (5) therefore imposes a cointegration restriction on the co-movements of the log-consumption ratio and log-price ratio series.

Equation (5) provides a structural equation that can be estimated consistently with cointegration techniques. From an economic perspective, this structural equation allows for a neat interpretation of θ where gross complementarity between private and government consumption corresponds to estimates of θ between zero and one, while estimates of θ greater than or equal 1 imply gross substitutability. It is important to point out that the estimation equation is void of the intertemporal elasticity of substitution. This allows us to focus on uncovering the intratemporal elasticity of substitution without having to make any stringent assumptions on the intertemporal elasticity of substitution.

It is worth noting that the theory does not provide an exclusive guidance on which variable, the consumption ratio or price ratio, to employ as the regressand. While our preferred specification selects the consumption ratio as the regressand, we also estimate the *IES* using the price ratio as the regressand. We go on to show in Sect. 5 that selecting the consumption ratio as the regressand as in Eq. (5) is indeed the preferred specification.

4 Econometric evidence

Following the results from the previous section, the basic equation of interest for our analysis of uncovering the *IES* between private and public consumption is given as:

$$\ln \left(\frac{C_t}{G_t} \right) = \theta \ln \left(\frac{P_t^s}{P_t^c} \right) + v_{it}. \tag{6}$$

Additionally, we specify the v_{it} to follow the process,

$$v_{it} = \alpha_i + \zeta_i' \mathbf{f}_t + u_{it}, \tag{7}$$

where α captures country-specific effects. Additionally, unlike previous studies that have estimated the *IES*, we explicitly allow for a set of unobserved common factors \mathbf{f}_t with country-specific ‘factor loadings’ ζ_i' to account for the unobserved factors and economic spillovers that may drive the relationship under consideration. This parameter indicates the impact of the factor on unit i , and u_{it} is a pure idiosyncratic error. The common factors by design do not only drive the consumption ratio, but also affects the price ratio. The latter is synonymous to arguments in Mundlak et al. (2008), Holly et al. (2010) and Eberhardt and Teal (2013). This generates a different type of endogeneity that is not easily remedied through instrumental variable estimations. Moreover, these common factors can encompass either weak factors, strong factors or both. The weak factors include local spillover effects that arise from shared cultural heritage, geographic proximity, economic and social interactions and integration (Banerjee and Carrion-i Silvestre 2017). On the other hand, strong factors capture more global factors such as global shocks or even synchronized changes in consumer preferences (e.g.,

financial crisis in 2008, the 1970s oil crisis, and the ongoing COVID-19 pandemic, which calls for almost synchronized policy actions across the world). Together, these common factors, weak and strong, should not be discounted as omitted variables but instead a set of latent drivers of these macroeconomic variables. In the presence of these common factors, one cannot correctly identify the parameter of interest θ unless the unobservable factors in the error term v_t are accounted for.¹³

Notice that if both $\ln(C_t/G_t)$ and $\ln(P_t^s/P_t^c)$ are difference-stationary I(1) processes, and v_t is a stationary I(0) process, then this implies that the two variables are cointegrated. We will provide formal evidence for this cointegration and unit root properties in Sect. 4.1. Here, the parameter θ can be estimated consistently from Eq. (6) even though there may be measurement errors or stationary omitted variables. That is, the gradient parameter can be estimated consistently without the assumption that the regressors are econometrically exogenous. This is possible because cointegration estimators possess super-consistency properties (Pedroni 2019).

Although Eq. (6) is the basic cointegrating equation employed by existing studies and is directly tied to the theoretical setup in section 5, a plethora of macro-econometric specifications arise to control for potential econometric issues. Specifically, an applied econometrician can either assume parameter homogeneity—in which case the key parameter θ is assumed to be the same across cross-sectional units (i.e., $\theta_i = \theta$) for all countries—or parameter heterogeneity across countries in the panel—in which case θ varies for each country. Here, we follow the literature and pool the data. Besides, pooling the data can lead to efficiency gains (Baltagi and Griffin 1997; Baltagi et al. 2008; Hsiao 2007). Importantly, we present homogeneity tests to formally confirm this assumption before proceeding with the estimation.

Moreover, the issue of cross-sectional dependence versus independence become an important assumption in estimating the IES under panel data. Furthermore, even if one assumes cross-sectional dependence, there are a variety of ways to deal with it depending on whether cross-sectional dependence in the error structure is weak or strong. For instance, does one apply cross-sectionally augmented means as in the Pesaran (2006) approach or should one address it by time demeaning the data as in Francois and Keinsley (2019) and Herzer and Morrissey (2013)? The choice of how cross-sectional dependence is treated can have a non-trivial impact on the accuracy of the estimated IES (see for example Sarafidis and Wansbeek 2012; De Hoyos and Sarafidis 2006, for a discussion on the treatment of cross-sectional dependence).

Furthermore, the choice of estimators is an important decision for the econometrician. In the context of nonstationary panel, there is a plethora of estimators—e.g., *panel dynamic OLS* (DOLS), *fully-modified OLS* (FMOLS) among other—that one can choose from. These estimators are designed to handle nonstationary data and are able to circumvent endogeneity issues arising from certain forms of simultaneity, omitted variables, measurement error, and reverse causality. This is because these estimators possess a superconsistency property under cointegration (see, Pedroni 2019). Nonetheless, each of them have their strengths and weaknesses. For instance, Kao and

¹³ Alternatively, one can identify θ if one has access to a valid and informative set of instruments, which is naturally challenging to come by. Specifically, it is challenging, and in some cases difficult, to find adequate and appropriate instruments using macroeconomic data is challenging because of complicated interrelationships (Temple 1999)

Chiang (2000) study the asymptotic distributions for the OLS, FMOLS, and DOLS, and finds that the DOLS outperforms both the OLS and FMOLS. These estimators do not traditionally address cross-sectional dependence. However, as we show in Sect. 5, they can be easily modified to address the issue of cross-sectional dependence.

In summary, with the exception of the assumption of pooling the data prior to estimation we do not make any additional assumptions on cross-sectional dependence, choice of estimator or regressand prior to estimation. Specifically, we remain agnostic and present an array of models that considers these assumptions individually or jointly. In the baseline specification in Eq. (6), we pool the data; hence, constraining the parameter θ to be common for all countries in the panel. We formally discuss and test for the assumption of homogeneity in Sect. 4.1.3. This assumption may seem stringent on face-value nonetheless, we discuss how the alternative of allowing heterogeneity in the estimated parameter can lead to severe misinterpretation or inaccurate inferences due to theoretical restrictions of θ . We then go on to estimate the IES under several sets of model assumptions and conduct a battery of post-estimation diagnostics to compare the validity of the estimated IES across models.

4.1 Data and pre-testing

We employ annual data for 1970 to 2018 from the World Development Indicators (World Bank, 2019) for 17 European economies. These countries were selected primarily due to data availability. Additionally, the start and end dates for the data are driven by missing observation for years leading up to 1970 and after 2018. The consumption ratio is derived by dividing household final consumption expenditure by general government final consumption expenditure, both in 2010 constant dollars. The corresponding prices are computed as the implicit price deflators, which are constructed by dividing the nominal private and government consumption series by their respective constant price series. Figure 3 depicts the two series for all 17 European countries in the sample. It is generally evident that there is strong persistency and co-movements of the two series. It is therefore intuitive, at least, to assume at face-value that the two series are individually $I(1)$ and potentially cointegrated. We now turn our attention to formally test these observations. However, for countries such as Luxembourg, Spain, and Switzerland, we observe a divergence in the movement of the two series. Moreover, the strength of co-movement varies across countries. This naturally prompts the need to go beyond pooled panel analysis and explore country-specific analysis.

Recall that preferences are stable, implying that the residual term v_{it} is stationary, and that our basic equation for estimation is a cointegrating regression if the regressor and regressand are both integrated processes. In this sense, the combination of stable preferences and the optimality conditions derived in Eq. (6) imposes a cointegration restriction on the co-movement of the log consumption and price ratio series. Since the assumption of stationarity is placed on the error term in the empirical model, the natural litmus test for cointegration regression in our case is to utilize residual-based cointegration tests. To this end, we formally test whether (i) the log consumption ratio and log price ratio are $I(1)$ processes, and (ii) whether the error term v_{it} is stationary

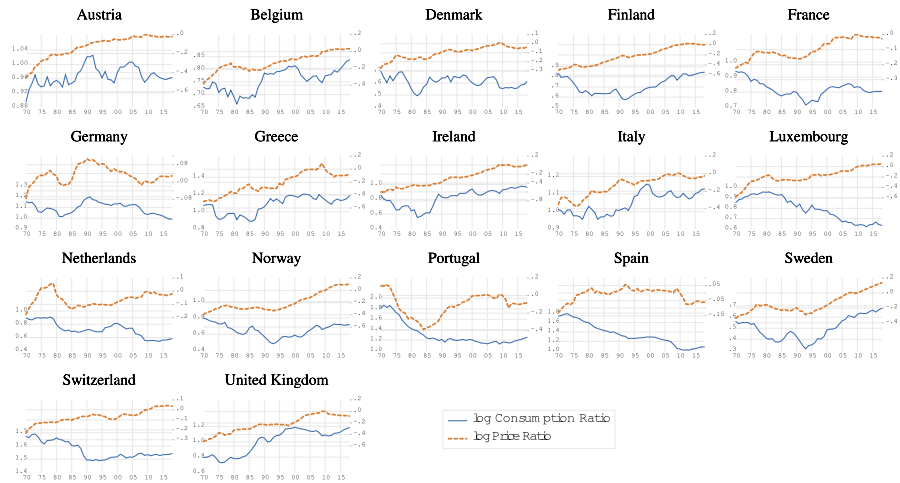


Fig. 3 Private and government consumption ratio and relative price. The horizontal axis is the time horizon (in years). The left and right y-axis represent the log consumption and relative price ratio, respectively

and whether there exists a cointegration relation in Eq. (6). In what follows, we conduct a thorough pre-testing analysis to confirm a cointegration relation in Eq. (6).

4.1.1 Unit root tests

It has been widely shown that most of the unit root tests for time series have low power and therefore accept the null of a unit root too often. The extension of unit root tests to a panel framework improves the power of unit root testing by incorporating information contained in the cross-sectional dimension. In this study, we pool the data for the 17 countries to perform four first generation panel unit root tests (Levin et al. (2002) (LLC), Breitung (2000), Im et al. (2003) (IPS), and Maddala and Wu (1999) (ADF)). The first two tests—LLC and Breitung—assume a common autoregressive coefficient across all cross sections, while the final two (IPS and ADF) allow for more flexibility by permitting the autoregressive coefficient to vary across cross-sections. These tests, however, do not account for cross-sectional dependence—they assume cross-sectional independence—which is inadequate and could lead to significant size distortions in the presence of neglected cross-sectional dependence (Baltagi and Pesaran 2007). Hence, in addition to these first generation tests we also consider the cross-sectionally augmented IPS (CIPS) test proposed by Pesaran (2007). The CIPS filters out any cross-sectional dependency by augmenting the ADF regression with the cross-section averages of lagged levels and first-differences of the individual series (See for example Herzer and Grimm 2012; Baltagi and Pesaran 2007, for a discussion on second-generation unit root tests).

Table 1 reports the formal panel unit root test results. It is evident from the table that all five tests fail dramatically to reject the unit root null hypothesis for the level series. However, the unit root null hypothesis is strongly rejected when we employ the first

Table 1 Unit Root Tests

Test Statistic	$\ln C_t^R$	$\ln P_t^R$	$\Delta \ln C_t^R$	$\Delta \ln P_t^R$
LLC	-1.0283 (0.1519)	-1.3144 (0.0944)	-7.827 (0.0000)	-9.284 (0.0000)
Breitung	0.7956 (0.7869)	0.3845 (0.6497)	-	-
IPS	-0.6203 (0.2675)	-1.5523 (0.0603)	-15.148 (0.0000)	-15.497 (0.0000)
ADF	37.073 (0.3291)	43.219 (0.1336)	280.989 (0.0000)	287.435 (0.0000)
CIPS	-2.331 (0.5160)	-2.276 (0.618)	-2.514 (0.0000)	-2.633 (0.0000)

Notes: p-values in round brackets. Probabilities for Fisher tests are computed using an asymptotic χ^2 -square distribution. All other tests assume asymptotic normality. For the level data, we allow for both individual country effects (c) and country-specific time trends (t). In the case of the first differenced data, allow for individual country effects (c). Automatic lag length selection with maximum of three lags were selected to adjust for autocorrelation. Andrews automatic bandwidth selection and Quadratic Spectral kernel were set for the exercise. The relevant 5% (1%) critical value for the CIPS statistics with is -2.71 (-2.85) with an intercept and a linear trend, and -2.20 (-2.36) with an intercept

differenced series. The results therefore confirm that the log price and consumption ratio series are non-stationary $I(1)$ processes.

4.2 Choosing the regressand

In the context of estimating the IES, existing studies often remain silent on the choice of regressand. Naturally, the IES can directly be recovered from the equilibrium relationship in Eq. (6) by setting the consumption ratio as the regressand (Dawood and Francois 2018; Kwan 2009). However, because the equilibrium condition offers no guide on which variable to set as the regressand, some studies such as Amano and Wirjanto (1998) instead use the price ratio as the regressand. Hence, these studies estimate the inverse of the IES and then recover the IES. In theory, irrespective of the choice of regressand, the econometrician should uncover the same IES. However, using an empirical example of bivariate models, Ng and Perron (1997) show that least-squares estimates can have very poor finite sample properties when normalized, with regards to choice of regressand, in one direction but are well behaved when normalized in the other. This occurs when one of the $I(1)$ variables is a weak random walk or is nearly stationary. In what follows, we provide discussions from the Ng and Perron rule, Granger causality tests from a panel vector error correction model, and pairwise Granger causality tests to provide some insight into selecting the regressand.

4.2.1 The Ng and Perron rule

As proposed by Ng and Perron (1997) and applied in Kwan (2009) and Dawood and Francois (2018), it is more desirable to put the more integrated series as the regressor (explanatory variable) and the less integrated series as the regressand (dependent variable). It is evident that most of the p -values obtained for the case of the level series suggests that the log consumption ratio has a stronger random walk component than the log price ratio (Table 1). That is, the log consumption ratio is less integrated than the log price ratio for the CIPS tests, but all the other tests suggest that the log price ratio is less integrated than the consumption ratio. The p -values from the unit roots tests suggest mixed results. In particular, while the CIPS test suggests estimating Eq. (6) with the consumption ratio as the regressand, all the other tests suggest employing the price ratio as the regressand. However, since the CIPS test controls for cross-sectional dependence, it is considered superior to the other tests. Thus, the Ng and Perron rule suggests using the consumption ratio as the regressand based on the CIPS test results.

4.2.2 Granger causality

We now turn our attention to utilizing Granger causality tests in guiding the direction of causality. We present two exercises: (1) pairwise Granger causality tests and (2) panel Granger causality based on a panel vector error correction model (VECM). We start with the standard pairwise Granger causality tests and report test results based on Dumitrescu and Hurlin (2012, henceforth DH) which controls for cross-sectional dependence and allows for heterogeneity of this causal relationship. Thus, for a given

Table 2 Pairwise Granger-Causality Tests

Null hypothesis	(1) F-Stat	(2) DH Tests
$\ln P_t^R$ does not Granger cause $\ln C_t^R$	1.766	3.466***
$\ln C_t^R$ does not Granger cause $\ln P_t^R$	7.878***	6.684***

Notes: *** and * indicate rejection of the null hypothesis that the variable X does not homogeneously cause the economic variable Y at the 1% and 10% significance level, respectively. The number of lags in the individual regressions is determined by the Schwarz information criteria with maximum lags set to 3. The p -values are computed using 100 bootstrap replications to account for the presence of cross-sectional dependence

Table 3 Panel Granger Causality based on Panel VECM

Dependent variable	Explanatory variable		Long-run ECT
	Short-run $\Delta \ln C_t^R$	$\Delta \ln P_t^R$	
$\Delta \ln C_t^R$	–	0.0837 (0.1069)	–0.002*** (–3.563)
$\Delta \ln P_t^R$	0.1012*** (0.005)	–	–0.003*** (–6.775)

Notes: ***, **, and * denote a rejection of the null hypothesis of no cointegration at the 1%, 5% and 10% levels, respectively. p values in round brackets. The number of optimal lags was established as two, using the Schwarz information criteria under the unrestricted panel VAR model. Short Run-The sum of the lagged coefficients for the respective short-run changes in the independent variable(s) are shown with their corresponding Wald χ -square and p values in brackets (). For the long-run, coefficients of the error correction term (ECT) are reported and t-statistics in brackets ()

pair of economic variables, X and Y , the null hypothesis of the DH-test is that X does not homogeneously cause Y . We also report results from the standard Granger causality tests.

Table 2 presents the results from the pairwise Granger causality tests. Columns 1 and 2 report the results from the standard Granger causality and DH tests, respectively. With the standard test, we find evidence of uni-directional causality going from the consumption ratio ($\ln C_t^R$) to the price ratio ($\ln P_t^R$). That is, the test rejects the null hypothesis that consumption ratio does not Granger cause price ratio at the 1% significance level while simultaneously failing to reject the null that the price ratio does not homogeneously cause consumption. Turning to the DH test, we reject the null hypothesis in both cases. Specifically, we reject the null that the price does not homogeneously cause consumption at the 1% significance level and vice-versa. This provides evidence of a bi-directional causal relationship.

Table 3 presents the results from the Granger causality based on a panel VECM, which test for both short-run and long-run causality. Guided by the Schwarz information criteria, the lag structure is set to two. The results in Table 3 show evidence of uni-directional short-run causality running from consumption to price. In the long-run,

however, there is evidence of bi-directional causal relationship between consumption and prices.

4.2.3 Discussion

The results from the Ng and Perron (1997) rule and the Granger causality tests provide a mixed conclusion on which variable to select as the regressand. While pre-estimation tests are useful in designing the empirical study, post-estimation diagnostics help increase validation and reliability of the estimates from the empirical design. Ng and Perron (1997) document that the choice of regressand has implications for residual-based unit-root tests for cointegration. Consequently, rather than keeping to one specification, we present empirical results from both scenarios and explicitly present post-estimation diagnostic tests to validate estimates of the IES. More precisely, we report detailed post-estimation diagnostic tests on the desirable features of well-behaved residuals along the lines of Eberhardt and Presbitero (2015). For completeness, we present cointegration tests with normalization in both direction.¹⁴

4.2.4 Cointegration tests

In this section, we present evidence of a cointegration relation in the main equation of interest, Eq. (6). The primary goal here is to validate the assumption that preferences are stable. This requires testing the stationarity property of the residuals. To this end, we test for the presence of cointegration in Eq. (6) using residual-based cointegration tests. Specifically, we employ four standard panel and group test statistics suggested by Pedroni (1999). The standard Pedroni tests, however, do not account for potential cross-sectional dependence. In the presence of cross-sectional dependence that may arise from multiple unobserved common factors, an assumption of cross-sectional independence can lead to biased inference (Herzer and Morrissey 2013; Baltagi and Pesaran 2007). In order account for cross-sectional dependence, we utilize the version of the standard Pedroni tests, which deals with cross-sectional dependence in the manner of Neal (2014). The strategy involves time demeaning of the data for each cross-sectional unit and variable (See, Neal 2014, for the theory and implementation details). For completeness, we report test results in which we assume cross-sectional independence.

In addition to the residual-based tests, we report cointegration tests based on Westerlund (2007). The cointegration test by Westerlund is based on structural rather than residual dynamics and therefore, do not impose any common factor restriction. Importantly, Westerlund (2007) compares small sample performance of the tests relative to the performance of the popular residual-based test by Pedroni (1999) and find good size accuracy, and that they are more powerful than the residual-based test.¹⁵ The test is designed to test the null by inferring whether the error correction term in a conditional

¹⁴ It is also worth noting that studies such as Amano and Wirjanto (1998), Kwan (2009) and Dawood and Francois (2018) do not account for possible change in estimation results depending on which variable is selected as the regressand.

¹⁵ I use the 'xtwest' command by Persyn and Westerlund (2008) in Stata for this exercise.

error correction model is equal to zero. If the null hypothesis of no error correction is rejected, then the null hypothesis of no cointegration is also rejected. Each test is able to accommodate individual-specific short-run dynamics including serially correlated error terms, non-strictly exogenous regressors, individual-specific intercept, and individual-specific slope parameters. We utilize bootstrap tests to account for cross-sectional dependence.

Table 4 presents results for two cases: one where the consumption ratio is used as the regressand and another scenario where the price ratio is used as the left-hand-side variable. Test results generally reject the null hypothesis of no cointegration at conventional levels of statistical significance. In particular, with the exception of the Group normalized statistic and the Pedroni Panel PP statistic, seven of the nine cointegration tests reject the null of no cointegration in the case where the consumption ratio is selected as the regressand. Similarly, six out of the nine tests in Table 4 rejects the null when the price ratio is employed as the regressand in Eq. (6). The non-rejection of the null hypothesis reinforces the need to estimate the model in Eq. (6) for the two choices of regressand.

5 Benchmark estimate of IES

In this section, we present the results from the panel cointegration regression. All regression specifications account for country-specific fixed effects. To explicitly highlight the relevance of how: (1) the choice of estimator, (2) assumptions on and treatment of cross-sectional (in)dependence, and (3) the choice of regressor and regressand in Eq. (6) impacts the size of the estimated IES. We report an array of estimators that consider the aforementioned scenarios. More specifically, for the choice of estimators we utilize the pooled versions of Dynamic-OLS (DOLS) estimator by Kao and Chiang (2000) and Mark and Sul (2003) and the Fully-Modified OLS by Pedroni (2001a). Without any modification, these estimators assume cross-sectional independence in the data. As previously discussed, in the presence of cross-sectional dependence, this can lead to inaccurate estimates. To account for potential cross-sectional dependence in the data, we employ two approaches. First, we use the cross-sectional demeaning approach applied in Herzer and Morrissey (2013) and Francois and Keinsley (2019). This approach subtracts the cross-sectional average of each variable in the original data employed in the estimation equation. Specifically, we replace each variable X_t in Eq. (6) by the transform \tilde{X}_t where,

$$\tilde{X}_t = X_t - \bar{X}_t, \quad \text{and} \quad \bar{X}_t = \frac{1}{N} \sum_{i=1}^N X_{it} \tag{8}$$

The estimators associated with the cross-sectionally demeaned (CD) variables are DOLS-CD and FMOLS-CD. It is important to note that while the cross-sectional demeaning approach has the advantage of avoiding over-parameterization as it preserves the number of regressors in the specification, it only addresses potential weak cross-sectional dependence. If the type of cross-sectional dependence is strong or

Table 4 Panel Cointegration Tests

Statistics	Econometrician's Choice of Regressand $\ln(C_t/G_t)$	$\ln(P_t^S/P_t^C)$
<i>Westerlund ECM panel cointegration tests^a</i>		
Panel normalized statistic (P_α)	-5.825**	-4.822
Panel statistic (P_τ)	-9.150***	-7.986*
Group normalized statistic (G_α)	-7.284	-5.769
Group statistic (G_τ)	-2.090***	-2.028*
<i>Engle-granger based cointegration tests^b</i>		
Kao (1999)	-3.637***	-0.726
Pedroni panel PP statistic	0.180	-2.789***
Pedroni panel ADF statistic	-1.560*	-2.289***
Pedroni group PP statistic	-1.330*	-2.438**
Pedroni group ADF statistic	-2.295**	-1.729**

Notes ***, **, and * denote a rejection of the null hypothesis of no cointegration at the 1%, 5% and 10% levels, respectively. ADF stands for augmented Dickey Fuller and PP stands for Phillips-Perron. The Kao (1999) test and all Panel statistic test the hypothesis that the panel is cointegrated as whole. In contrast, Group statistics test the alternative that there is at least one individual that is cointegrated. All tests are implemented with a constant in the test regression. The null hypothesis here is no cointegration

^a We assume cross-sectional dependence in data and hence, employ the bootstrap approach. The tests were carried under robust p -values, which are based on the bootstrapped distribution. The bootstrap test are implemented using 100 bootstraps replications. The bandwidth and number of lags are set to the rule $\text{int}[4(T/100)^{2/9}]$ where $T = 49$. Additionally, for semi-parametric corrections, the Bartlett kernel is employed

^b All test statistics are distributed $N(0,1)$, under a null of no cointegration. Cross-sectional independence is assumed. The Pedroni tests the number of lags was determined by the Akaike criterion with a maximum of four lags. We assume Newey-West fixed bandwidth, and for semi-parametric corrections we employ the Quadratic Spectral kernel

Table 5 Estimated Values of IES with Consumption Ratio as Regressand

Panel A	Independence			Dependence	
	[1] DOLS	[2] FMOLS	[3] DOLS-CD	[4] FMOLS-CD	[5] CCEP
$\hat{\theta}$	0.311*** (0.058)	0.297*** (0.061)	0.738*** (0.077)	0.901*** (0.074)	0.515*** 0.042
Panel B: Diagnostics					
\hat{u}_t : CD <i>p</i> -value	0.002	0.000	0.444	0.003	0.000
$\hat{\varepsilon}_t$: <i>I</i> (1) <i>p</i> -value (ADF)	0.007	0.003	0.068	0.048	0.019
\hat{u}_t : <i>I</i> (1) <i>p</i> -value (CIPS)	0.560	0.265	0.034	0.296	0.371
RMSE	0.086	0.107	0.069	0.085	0.050
Observation	743	816	744	816	833

Notes: Standard error in round brackets. *** represents statistical significance at the 1% level. All estimation equations allow for country-specific effects. Columns (1)–(2) assume cross-sectional independence in the data while Columns (3)–(5) assume cross-sectional dependence in the data

more complex, this strategy of accounting for cross-sectional dependence would not be sufficient. Specifically, as discussed in De Hoyos and Sarafidis (2006), while cross-sectional (time-) demeaning removes the mean impact of the factors, in the polar case where the variance of the coefficient on the factor loadings λ_i in Eq. (7) grows large, time demeaning will be less effective. This is because even if the mean impact of the factors has been removed, there will still be a considerable amount of cross-sectional dependence left out in the disturbance (see, De Hoyos and Sarafidis 2006; Sarafidis and Wansbeek 2012, for detailed discussion).

An alternative and more powerful approach to address cross-sectional dependence is to include the cross-sectional averages of all the variables in the regression equation á la Pesaran (2006) and Holly et al. (2010). This technique is a straightforward way of dealing with multi-factor and more complex cross-sectional dependence in the data. Consequently, we employ the pooled common correlated effect (CCEP) estimator by Pesaran (2006), which is naturally designed to deal with cross-sectional dependence in the data head on.¹⁶

Finally, recall that economic theory in Sect. 4 suggests that there is no “silver bullet” on which variable to employ as regressand (or regressor). Moreover, using the Ng and Perron rule of thumb does not offer a clear decision on which variable, $\ln(C_t/G_t)$ or $\ln(P_t^g/P_t^c)$, to employ as the regressand or regressor. To this end, we report results for a case where the consumption ratio is employed as the regressand (e.g., Dawood and Francois 2018; Kwan 2009) and another scenario where the price ratio is used as the regressand (e.g., Amano and Wirjanto 1998). In summary, we run 10 different regressions to uncover the intratemporal elasticity of substitution. More importantly, we report several post-estimation results to compare the validity of estimates from the estimation choices.

Tables 5 and 6 present the main results. We begin by focusing on the results from Table 5, which utilizes the consumption ratio as the regressand. Panel A presents the

¹⁶ See Holly et al. (2010) for details.

pooled estimates of the IES. There are a number of observations: First, all estimated values of θ are positive, less than 1, and range between a low value of 0.297 to a high value of 0.901. The estimates are statistically significant at conventional levels. The positive estimates satisfy the preference properties of non-negativity of the IES. Additionally, because all the estimated values of θ are below unity, they suggest that private and public consumption are *gross complements* in these European economies. The finding of gross complementarity is similar to findings in Dawood and Francois (2018) and Kwan (2009) in the case of African and East Asian countries, respectively. The results are, however, in contrast to findings in Amano and Wirjanto (1998) who estimate the IES to be 1.56 in the case of the United States, implying *gross substitution* between the two goods. Second, estimates from the FMOLS and DOLS estimators that assume cross-sectional independence yield the smallest estimates, 0.297 and 0.311, respectively. In contrast, accounting for the presence of cross-sectional dependence of any form, as shown in columns (3)–(5) drastically increases the size of the estimated IES. In particular, the FMOLS-CD produces the largest θ amongst the estimates that accommodate cross-sectional dependence. The CCEP estimator on the other hand yields the smallest value of 0.515 amongst the estimators that control for cross-sectional dependence. Finally, the DOLS-CD uncovers an IES of 0.738. This generally suggests that an assumption of cross-sectional dependence (or independence) can affect the size of the estimated IES non-trivially.

To better appreciate these estimates, we present a number of post-estimation diagnostic tests that focus on the behavior of the residuals from the estimators in Table 5. The diagnostics are meant to evaluate the performance and efficiency gains of the selected estimators. More importantly, they provide a yardstick for validating a particular model assumption *ex post*. The diagnostic tests are presented in Panel B of Table 5. The post-estimation diagnostics include cross-sectional dependence given by the Pesaran (2015) cross-sectional dependence test, unit roots (CIPS) tests, and the root mean square error (RMSE). In the context of cross-sectional dependence tests, a desirable property of the residuals from the estimators should be that they exhibit cross-sectional independence. Hence, a failure to reject the null hypothesis of cross-sectional independence is the desired outcome. Evidently, from Panel B, the null of cross-sectional independence is rejected for all estimators except for the DOLS-CD estimator. Specifically, the latter estimator fails to reject the null of cross-sectional independence in the error structure. Turning to the unit root test, we apply two unit root tests—the Maddala and Wu ADF and the CIPS unit root tests—to the residuals from the regressions in Table 5. The goal is to check the stationarity property of these residuals. Recall that a failure to reject the null of unit root violates the stationarity property of preferences described in Sect. 4. As shown in Panel B in Table 5, the p -values associated with the ADF unit root test decisively suggest the rejection of the null of unit root at the 10% significance level or better. However, under the CIPS unit root test, the null is only rejected in the case of DOLS-CD. Finally, the RMSE ranks the CCEP followed by the DOLS-CD as the best predictor of the observed data as

Table 6 Estimated Values of IES with Price Ratio as Regressor

Panel A	Independence			Dependence	
	[1] DOLS	[2] FMOLS	[3] DOLS-CD	[4] FMOLS-CD	[5] CCEP
$\hat{\beta}$	0.330** (0.068)	0.345** (0.066)	0.412** (0.038)	0.444** (0.037)	0.318** (0.026)
Implied ($\theta = 1/\beta$)	3.030	2.899	2.427	2.252	3.145
Panel B: diagnostics					
$\hat{\varepsilon}_t$: CD p -value	0.000	0.000	0.004	0.000	0.000
$\hat{\varepsilon}_t$: $I(1)$ p -value (ADF)	0.609	0.582	0.892	0.802	0.004
$\hat{\varepsilon}_t$: $I(1)$ p -value (CIPS)	0.192	0.208	0.952	0.969	0.139
RMSE	0.095	0.113	0.056	0.061	0.039
No. of Observation	749	816	781	816	833

Standard error in round brackets. ** represents statistical significance at the 1% level. All estimation equations allow for country-specific effects. Columns (1)–(2) assume cross-sectional independence in the data while Columns (3)–(5) assume cross-sectional dependence in the data

they have the smallest RMSE, respectively.¹⁷ Overall, the DOLS-CD model which accounts for cross-sectional dependence outperforms the other competing models.

We now turn our focus to Table 6. The results presented here are from the estimations where the price ratio is employed as the regressand. We employ the same set of estimators from Table 5. It is important to note that with the price ratio as the regressand, the estimated parameter, β is the inverse of the IES. Hence, to recover the IES, one needs to take the inverse of the estimated values. From Table 6, β is estimated to lie between 0.318 and 0.444, and these estimates are statistically significant at the 1% level. This implies that the intra-period elasticity of substitution parameter, θ , ranges from 2.252 to 3.145, suggesting that private and public consumption are *gross substitutes*. These results are in stark contrast to the *gross complements* finding between the two goods when the consumption ratio is employed as the regressand. Interestingly, the post-estimation diagnostics presented in Panel B of Table 6 generally suggest that the residual from these regressions violate the set of desirable properties discussed earlier. In particular, compared to the behavior of the residuals from the estimations in Table 5, the null hypothesis of cross-sectional independence of the residuals are strongly rejected for all estimations. Additionally, test statistics from the unit root test show that the residuals from all the estimations do not poses the required stationarity property implying that preferences are not stable. Finally, while the RMSE for the estimations in Table 6 is on average lower (i.e., 0.0728) than in the case where the consumption ratio is employed as the regressand (i.e., 0.0794), the difference is marginal. In summary, one can confidently conclude that the results in Table 5 are the more reliable estimates, and the DOLS-CD estimator is preferred over the other esti-

¹⁷ Note that a small RMSE generally implies our model is good at predicting the observed data, whereas if RMSE is large, this generally means our model is relatively fails to account for important features underlying our data.

mators. Consequently, the preferred estimated IES is 0.738 as given by the DOLS-CD estimator in column 3 of Table 5.

5.1 Cross-sectional augmented distributed lag (CS-DL) estimator

It is worth mentioning that failing to account for cross-sectional dependence would bias parameter estimates only if unobservable factors are correlated with regressors (as stated in the paper), but they would reduce parameter efficiency if these factors are correlated with the dependent variable. To purge the effect of strong CD from estimates, we adopt the pooled Common Correlated Effect estimator (CCEP), developed by Pesaran (2006), as an alternative to cointegration regression techniques, such as DOLS and FM-OLS run on cross-sectional de-meaned data. Unfortunately, the CCEP is not the best candidate to this aim as it is a static procedure of regression and hence is incapable of purging the effect of dynamic adjustment and simultaneous feedback between regressand and regressors (unlike DOLS and FM-OLS). The set of dynamic procedures, capable of estimating a potentially cointegrated regressions and account for strong CD, includes the cross-sectionally augmented version of the Auto-Regressive Distributed Lag (CS-ARDL) or the Cross-Sectionally augmented Distributed Lag (CS-DL) model by Chudik et al. (2017, 2013).

These two estimators have their merits and drawbacks as discussed by Chudik et al. (2013). In particular, the main advantage of the CS-DL approach relative to the CS-ARDL approach is its superior small sample performance when the time series dimension of the panel is moderate. Specifically, for the consistency of the ARDL estimates, sufficiently long lags are necessary, whereas specifying longer lags than necessary can lead to estimates with poor small sample properties. The CS-DL method is more generally applicable and requires only that a truncation lag order be selected. A drawback of the CS-DL technique relative to the CS-ARDL approach is that the CS-DL estimates of long-run effects are not consistent when there is significant feedback from the regressand to regressor. Nonetheless, Chudik et al. (2016) argue that even with this bias, the performance of CS-DL in terms of RMSE is much better than that of the CS-ARDL approach when T is moderate (which is the case in our empirical application). Furthermore, the CS-DL approach is robust to a number of departures from the baseline specification, such as residual serial correlation, and possible breaks in the error processes. To this end, we employ the CS-DL estimator, which suits our purposes given the small-sample time series properties of our data (i.e., $T = 49 < 100$) and our small size of the cross-sectional units $N = 17$. We augment the model with four lags of the cross-sectional averages of the dependent variable. The lag length for the cross-sectional average is selected using the rule of thumb of $T^{1/3}$ suggested by Chudik et al. (2017, 2013).¹⁸

Table 7 reports the results from the CS-DL estimation. For completeness, we report the estimates of the IES for different specifications. While the estimation with the consumption ratio as the dependent variable is our preferred model, we report the results from the case where the price ratio is employed as the dependent variable (columns 5–

¹⁸ See also Eberhardt and Presbitero (2015) for additional discussion on the selection of the lag length for the cross-sectional averages.

Table 7 CS-DL Estimates of the IES

	Dependent variable: C_t/G_t				Dependent variable: P_t^s/P_t^c			
	(1) CS(1 4)	(2) CS(2 4)	(3) CS(3 4)	(4) CS(4 4)	(5) CS(1 4)	(6) CS(2 4)	(7) CS(3 4)	(8) CS(4 4)
Panel A: estimated IES								
Coefficient Estimate	0.546*** (0.120)	0.575*** (0.121)	0.616*** (0.123)	0.665*** (0.131)	0.316** (0.132)	0.321*** (0.137)	0.323*** (0.142)	0.326*** (0.144)
Panel B: diagnostics								
$\hat{\epsilon}_t$: CD p -value	0.001	0.0203	0.233	0.940	0.000	0.001	0.0126	0.120
$\hat{\epsilon}_t$: $I(1)$ p -value (ADF)	0.000	0.004	0.000	0.000	0.000	0.084	0.037	0.005
$\hat{\epsilon}_t$: $I(1)$ p -value (CIPS)	0.000	0.002	0.000	0.000	0.006	0.152	0.000	0.000
RMSE	0.0496	0.0483	0.0468	0.0452	0.0363	0.0352	0.0339	0.0324
No. of observation	765	765	765	765	765	765	765	765

The dependent variable is the change in consumption ratio. ***, **, and * indicate significance at the 1, 5, and 10% level. P-values in parenthesis. The estimation is the cross-section augmented distributed lag (CS-DL) by Chudik et al. (2013). The specific regression model is given as $C_{it}/G_{it} = c_i + \theta P_{it}^s/P_{it}^c + \sum_{i=0}^{p-1} \delta' \Delta C_{it}/G_{it,t-i} + \omega_y \overline{C_{it}/G_{it}} + \sum_{i=0}^L \omega_{x,i} \Delta P_{it}^s/P_{it}^c + e_{it}$

8) to uncover the IES to highlight the potential model misspecification. The estimates from columns 1–4, where the consumption ratio is employed as the regressand, reveal an IES value ranging from 0.55–0.67. Importantly, the post-estimation diagnostics for the presence of unit root in the residual is strongly rejected for all specifications. While the null of CD test is rejected in columns 1 and 2, we fail to reject the null for columns 3 and 4. This suggests that the specifications in columns 3 and 4 produce a more reliable estimates of the IES (0.62 and 0.67, respectively), which are in line with estimates from the preferred DOLS results in Table 5. Switching to the case where the price ratio is used as the dependent variable á la Amano and Wirjanto (1998) (columns 5–8), the results uncover unusually large estimates of the IES (ranging from 3.16 to 3.07).¹⁹ Notice that these values, when embedded in the standard RBC model described in Section 2, imply a large negative response of private consumption, which fails to predict the empirical evidence of a positive consumption response to an increase in government consumption. Importantly, only the specification in column 8 satisfies all the post-estimation checks. These further suggest that estimates using the price ratio will lead to misspecification and implausible estimates of the IES.

What does the preferred estimate imply for Edgeworth complementarity (substitutability) between private and public consumption? As discussed in Sect. 2, under a CRRA utility function, the sign of the cross-partial derivative, $U_{CG} = \partial(\partial u/\partial C)/\partial G$ governs the Edgeworth substitutability (complementarity) between the two goods. Specifically, given the utility function in this study, Amano and Wirjanto (1998) show that the cross partial derivative, U_{CG} depends on the difference between the intertemporal elasticity of substitution ($1/\gamma$) and the IES (i.e., $sgn[U_{CG}] = sgn[1/\gamma - \theta]$). Here, private and public consumption are therefore Edgeworth complements (substitutes) if the intertemporal elasticity of substitution is greater (less) than the intratemporal elasticity of substitution. If the two preference parameters are equal, then changes in government consumption have no impact on the marginal utility of private consumption. Fixing the intertemporal elasticity at 0.8 as in Havránek (2015), one can observe that $sgn[U_{CG}] > 0$ for the preferred estimate of the IES (i.e., 0.738) as given by DOLS-CD estimate in Table 5 as well as the estimates from columns 3 and 4 from the CS-DL estimates in Table 7. The positive sign of the cross partial derivative, U_{CG} , suggests that private and government consumption are Edgeworth complements in the Pareto sense—a result consistent with Fiorito and Kollintzas (2004). The finding that the two goods are Edgeworth complements is robust to lower assumed values of the intertemporal elasticity up to a minimum value of 0.79.

The immediate implication for policy is that with Edgeworth complementarity between the two goods, an increase in government consumption would increase the marginal utility of private consumption. If the rise in the marginal utility is stronger than the standard wealth effect induced by tax- (or deficit-) financing of the increase in government consumption, then private consumption would rise. Indeed, as demonstrated in Fig. 2 the baseline estimate of θ of 0.738 implies an increase in private consumption following a rise in government consumption. Consequently, a fiscal expansion that increases government consumption would have Keynesian effects on real output

¹⁹ For the case where the price ratio is used as the regressand, the IES is given by $\theta = \frac{1}{\beta}$, where β is the coefficient estimate.

in European economies. These Keynesian effects are similar to finding in Amendola et al. (2020).

6 Heterogeneity in estimated IES

A particular feature of the empirical model in the previous section is the assumption of homogeneity. This assumption leads to efficiency gains from pooling the data, which is a desirable feature for the reliability of estimates (see, Baltagi and Griffin 1997; Baltagi et al. 2008, for a discussion). Moreover, one can generate more accurate predictions for individual outcomes by pooling the data rather than generating predictions of individual outcomes using the data on the individual in question. Specifically, if individual behaviors are similar, conditional on certain variables, panel data provide the possibility of learning an individual's behavior by observing the behavior of others (Hsiao 2007). Thus, it is possible to obtain a more accurate description of an individual's behavior by supplementing observations of the individual in question with data on other individuals.

In this study, however, we employ aggregate consumption data at the country-level. Additionally, the prior assumption of pooling the data induces an additional level of aggregation, which while it has its benefits, inadvertently over-aggregates the data. This invokes a strong "representative agent" assumption. Here, if individual countries in the panel are heterogeneous in terms of the size (and even the sign) of the IES, the time series properties of the aggregate data would be starkly different from those of dis-aggregate data, which in this case is the country-level data (Granger 1988; Pesaran 2003; Pedroni 2001b; Chudik et al. 2017). Importantly, policy evaluation based on aggregate data may be grossly misleading given the latter. These motivate the question: *How large is cross-country variation in preferences relative to the pooled estimate?* To answer this question, we relax the assumption of homogeneity of the IES and allow it to vary across countries. We therefore employ the Pesaran (2006) Common Correlated Effects Mean Group (GM-CCE) estimator, which allows for the IES to vary across each country and accounts for cross-sectional dependence in the data.²⁰ It is worth mentioning and acknowledging that, since consistency of group mean estimates is ensured by the large number of cross-sectional units, the GM-CCE estimates should be taken with caution. Given the limited number of countries used in the regression (i.e., $N = 17$), we follow Bond et al. (2010) and also report robust mean of single-country elasticities.²¹ Finally, we employ the consumption ratio as the dependent variable.

Table 8 presents the results from the heterogeneous panel estimation. The last row reports the panel group-mean estimate of the IES, which uncovers an estimated value of θ equal to 0.59.²² This estimated value is smaller in size compared to the desired pooled estimate from DOLS-CD estimator in Table 5 (i.e., 0.738). Nonetheless, the

²⁰ We use the Stata command `xtmg` to implement the Common Correlated Effects Mean Group estimator (see, Eberhardt 2012).

²¹ We are grateful to one of our anonymous referees for this suggestion.

²² The panel group-mean estimator (θ^M) is computed as $\hat{\theta}^M = \frac{1}{N} \sum_{i=1}^N \theta_i$, where θ^i is the country-specific estimate for the i th country in the panel and $N = 17$ is the number of countries in the panel.

Table 8 Country-specific estimates of IES

Country	CCE Estimator		Robust regression	
	EIS, $\hat{\theta}_i$	Std. Error	EIS, $\hat{\theta}_i$	Std. Error
Austria	0.3820***	(0.0841)	0.3954***	(0.0893)
Belgium	0.0097	(0.1637)	0.0103	(0.1860)
Denmark	1.169***	(0.2205)	1.1647***	(0.2321)
Finland	0.3965**	(0.1905)	0.4372**	(0.2146)
France	0.8224***	(0.1737)	0.7378***	(0.1619)
Germany	0.8832***	(0.1579)	0.9329***	(0.1666)
Greece	0.3746**	(0.1546)	0.4990***	(0.1524)
Ireland	1.2360***	(0.2488)	1.2654***	(0.2560)
Italy	0.3034***	(0.0882)	0.1953**	(0.0722)
Luxembourg	0.5642***	(0.0892)	0.5101***	(0.0884)
Netherlands	0.4077**	(0.1615)	0.4520***	(0.1752)
Norway	0.8850***	(0.0672)	0.8776***	(0.0703)
Portugal	-0.0414	(0.1844)	-0.0694	(0.2037)
Spain	0.6759***	(0.2555)	0.6344**	(0.2744)
Sweden	0.7883***	(0.0905)	0.8205***	(0.0979)
Switzerland	0.0972	(0.2750)	0.2848	(0.3079)
United Kingdom	0.9986**	(0.4418)	1.1346**	(0.4784)
Mean Group estimator (GM-CCE)	0.5854***	(0.0942)		

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The robust regression is robust estimate of robust mean of single-country elasticities. We use the `rreg` command in Stata. Each country regression is augmented with contemporaneous cross-sectional mean of the regressor and regressand to control for global factors that may be omitted in country regressions

group-mean estimate still implies that on average, private and public consumption are *gross complements*. More importantly, when combined with the relevant intertemporal elasticity of substitution, this estimated value suggests that the two consumption goods are *complements* in the *Edgeworth-Pareto* sense. Consequently, the conclusion from the heterogeneous panel estimator reinforces the main inference from the baseline results produced by the homogeneous case. Second, the table reports country-specific estimates of the IES, which highlights strong heterogeneity across countries. The estimates from both the CCE country-by-country results and the robust regression are similar. The results can be categorized into three groups: (1) Countries where the IES is positive, less than unity, and statistically significant. The countries include: Austria, Finland, France, Germany, Greece, Italy, Luxembourg, Netherlands, Norway, Spain, Sweden, and United Kingdom although the United Kingdom is closer to 1. In these economies, private and government consumption *gross complements*. (2) Countries where the estimated IES is positive, greater than 1, and statistically significant. Here, there are two countries that fall in this category—i.e., Denmark and Ireland. In these countries, the two goods in question are best classified as *gross substitutes*. (3) Economies including Belgium, Portugal, and Switzerland where the estimated IES is

statistically insignificant, suggesting that the null hypothesis that θ is equal to zero cannot be rejected. As discussed in Dawood and Francois (2018), this implies that either private and government consumption are perfect complements or that the identifying assumptions in this study do not hold for these countries. The heterogeneous analysis sheds light on why the panel estimates consistently found private and public consumption to be gross complements.

Similar to the panel analysis, we can determine whether the “well-defined” country-by-country estimates imply Edgeworth substitutability/complementarity in individual countries. More precisely, setting the intertemporal elasticity of substitution ($1/\gamma$) to a reasonable value of 0.8 as in Havránek (2015), we are able to demonstrate the importance and subsequent implication of the size of the IES across countries. With $1/\gamma$ fixed to 0.8, it is evident that private and public consumption are *Edgeworth complements* in all the countries where the IES is less than unity except for Denmark and Ireland. That is, the sign of the cross partial U_{CG} is greater than zero for these countries but for Germany, Greece, Norway, and United Kingdom where it less than zero (i.e., sign of $U_{CG} = [0.8 - 0.8224] < 0$, $U_{CG} = [0.8 - 0.8832] < 0$, $U_{CG} = [0.8 - 0.885] < 0$ and $U_{CG} = [0.8 - 0.9986] < 0$, respectively). For the two countries where the IES is greater than one, private and public consumption are unambiguously *Edgeworth substitutes*. The conclusions from the country-by-country CCE estimator hold true for the results from the robust regressions. These findings highlight the fact that the same fiscal policy involving changes in government consumption will likely yield very different outcomes in different countries, implying that policy design should be country-specific.

6.1 Explaining the cross-country heterogeneity

The heterogeneity analysis in Table 8 reveals that the size of the IES varies largely across countries. This pushes for the need to understand the factors that may explain the observed variation in the estimated IES. The natural correlates to consider are the components of government consumption—i.e., public spending on defense, education, health, and public order and safety. According to the functional definition of government consumption, these components can be classified into pure public goods (i.e., defense spending, law courts, and public order and safety) and merit goods (i.e., education and health). *A priori*, one would expect that the pure public good components would be negatively associated with the size of IES. This is because these components do not have immediate substitutes that can be provided by the private sector, at least on a large scale (Evans and Karras 1996). This therefore reduces the substitutability of these public goods. In contrast, merit goods such as public education and health can be provided by the private sector; hence, conditional on quality, they are likely to be easily substitutable with private education and health, respectively. Consequently, one would expect a positive relationship between these merit goods and the size of the IES. Nonetheless, these inherent characteristics of public and merit goods may not fully summarize the plausible relationship these goods may have with the IES. More precisely, merit goods are thought to have strong positive externalities, and are therefore complementary to private consumption. Thus, one can also expect a negative

relationship between merit goods and the degree substitutability as measured by the size of the IES. This presents a more complex relationship between the IES and the components of public consumption.

Beyond these aforementioned primary correlates, the size of government, which is given as government consumption as a share of GDP, can be a predictor of the IES. Here, as governments get bigger, they are likely to start providing more merit goods and services such as education and health relative to the pure public good/services such as defense they traditionally provide (Karras 1994). In this sense, since merit goods can be provided by the private and are therefore more substitutable for private consumption, one would expect that the size of the IES will be increasing in government size as government provides more merit goods.

Table 9 presents the OLS regression estimates for how the size of the IES is related to the share of defense, public order and safety, law court education, health spending in total expenditure, and government size. We break our analysis into several parts by examining the association between the merit and public components of public consumption and the IES (Columns 1 and 2, respectively). We then combine all the predictors and examine how each component is associated with the IES (Column 3). Finally, Column 4 presents the full model specification that includes all the components from columns 1 and 2, government size, and its squared value as predictors of the size of the IES.

We start our discussion with Column 1 in Table 9. The column documents that the merit good components—i.e., education and health—varies positively with the size of the estimated IES. Switching to Column 2, Table 9 indicates that there is no statistically significant relationship between the estimated IES and pure public goods components, which comprise defense spending, law courts, and public order and safety spending in total expenditure. Furthermore, Columns 3, which includes both public and merit goods as potential correlates in the regression, reinforces the results from Column 1. Finally, Column 4, which represents the full model, still shows that health expenditure is still positively correlated with the IES; however, the coefficient of education is negative but not statistically significant. Interestingly, the coefficient of public order, which has been negative and not statistically significant in the other specification, is now negative and statistically significant at the 10% level. Additionally, we find a U-shape relationship between the size of IES and government size. This finding suggests that there is a point beyond which as governments gets bigger, the degree of substitutability between government and private consumption becomes stronger.

In summary, there is some evidence that the variation in the IES is positively correlated with the share of health, and to some degree education, expenditures in government consumption and the government size. On the other hand, we find that public order and safety has a negative relationship with the IES, suggesting that an increase in the public good public order in government consumption will likely strengthen the complementarity between public and private consumption at the aggregate level. These findings are analogous to findings in Evans and Karras (1996), who find that non-education component of non-defense spending such as health drives the substitutability between private and government consumption. Meanwhile, similar to public order, the authors find that the higher share of defense spending in government expenditures, the stronger the complementarity between private and public consumption.

Table 9 Correlates of the IES at the country level

	Dependent variable: estimated IES (θ)			
	(1) Merit goods	(2) Public goods	(3) Merit + public components	(4) Full Model
Education (% total expenditure)	0.062** (0.028)	-	0.071** (0.029)	-0.031 (0.042)
Health (% total expenditure)	0.105*** (0.017)	-	0.105*** (0.025)	0.237*** (0.066)
Defense (% total expenditure)	-	0.025 (0.089)	0.018 (0.055)	0.064 (0.063)
Public order and safety (% total expenditure)	-	-0.168 (0.193)	-0.173 (0.138)	-0.260* (0.131)
Law courts (% total expenditure)	-	0.617 (0.669)	0.727 (0.548)	0.836 (0.459)
Government size	-	-	-	-0.957*** (0.333)
Government size (squared)	-	-	-	0.024** (0.008)
R^2	0.460	0.065	0.524	0.740

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Robust standard errors in parentheses. Number of observation is 16. Luxembourg was excluded from the regression analysis because of the negative sign of the estimated IES. The p -value for the joint significance of government size and its squared is 0.049

7 Concluding remarks

Accounting for utility-enhancing public consumption is important in mediating the effect of changes in government consumption on private consumption. In this paper, we combine theory and empirical work to uncover the intratemporal elasticity of substitution between private and public consumption in European economies. The empirical work explicitly accounts for cross-sectional dependence that may arise due to global shocks and economic spillovers. The latter makes this study the first in the literature to address the issue of cross-sectional dependence while estimating the IES in the context of panel data. Importantly, we remain flexible by adopting several estimators and data treatment in the empirical study. We then rely on simple but effective post-estimation diagnostics to guide the validity of the estimated IES. We find point estimates of the intratemporal elasticity of substitution that reveal that for the plausible values of the corresponding intertemporal elasticity of substitution, government and private consumption are best described as *Edgeworth complements* in European economies. The results imply that an increase in government consumption increases the marginal utility of private consumption, which offsets the negative wealth effect induced by financing the increase in public consumption. Stated differently, an increase in government consumption can induce Keynesian effects through a positive marginal utility of private consumption. In contrast, fiscal consolidation that cut government consumption can adversely impact output through a negative marginal utility of private consumption. This last result is similar to finding in (Barrell et al. 2013).

Policy-wise, weaker economic growth in European economies compounded with the economic impact of the ongoing COVID-19 pandemic calls for a strong need for fiscal stimulus in these economies. These findings suggest that fiscal expansions that increase government consumption can stimulate aggregate demand via a marginal utility channel of private consumption. The results further reinforce recent arguments that fiscal stimulus packages that comprise large government consumption components may be effective at stimulating aggregate demand (see, Boehm 2019, for example).

A The simple RBC model

The model is a standard RBC model with utility-enhancing government consumption and capital adjustment costs. The instantaneous utility function for households is given by $u(C_t^e, n_t) = \frac{(C_t^e)^{1-1/\gamma}}{1-1/\gamma} + \psi \ln(1 - n_t)$, where ψ is the level parameter that determines the relative importance of leisure in the utility function, $C_t^e = (\lambda C^{1-1/\theta} + (1 - \lambda)G^{1-1/\theta})^{1/(1-1/\theta)}$ is effective consumption, C_t is private consumption, G_t is public consumption and n_t is the number of hours worked. The representative agent finances consumption (C_t) and investment (i_t) decision with her earnings from labor income $w_t n_t$ and capital income $r_t k_t$ where w_t and r_t are real wages and real rental rate of capital (k_t), respectively. The shadow value of wealth associated with the household budget constraint is μ_t . Output (y_t) in the economy is produced according to the Cobb–Douglas production technology $y_t = A k_t^\alpha n_t^{1-\alpha}$ where A is the constant level total factor productivity (TFP). The government sector finances its consumption

G_t , which follows an AR(1) process, with lump-sum taxes, T_t . The full model is described in Eq. (A.1)–(A.10), and the parameters along with their corresponding values are provided in Table 10.

$$\mu_t = \lambda C^{-1/\theta} (C_t^e)^{1/\theta - 1/\gamma} \tag{A.1}$$

$$\mu_t = \psi / (w_t(1 - n_t)) \tag{A.2}$$

$$\mu_t = \beta E_t(\mu_{t+1}(1 + r_{t+1} - \delta + \kappa(i_{t+1}/k_t - \delta) + \frac{\kappa}{2}(i_{t+1}/k_t - \delta)^2)) / (1 + \kappa(i_t/k_{t-1} - \delta)) \tag{A.3}$$

$$y_t = Ak_t^\alpha n_t^{1-\alpha} \tag{A.4}$$

$$w_t = (1 - \alpha)y_t/n_t \tag{A.5}$$

$$r_t = \alpha y_t/k_t \tag{A.6}$$

$$G_t = T_t \tag{A.7}$$

$$k_t = (1 - \delta)k_{t-1} + i_t \tag{A.8}$$

$$\log(G_t) = (1 - \rho_g) \log(\bar{G}) + \rho_g \log(G_{t-1}) + \epsilon_g \tag{A.9}$$

$$y_t = C_t + i_t + G_t + \frac{\kappa}{2}(i_t/k_{t-1} - \delta)^2 \tag{A.10}$$

B Calibration of model parameters

Table 10 Standard Calibration of Parameters and Steady States

Description	Parameter	Calibration
Discount factor	β	0.995
Elasticity of sub. between C and G	θ	0.738
Intertemporal elasticity of substitution	$1/\gamma$	0.8
Cost share of capital	α	0.36
Capital depreciation rate	δ	0.01
AR(1) coefficient in government spending process	ρ_g	0.8
Capital adjustment cost parameter	κ	0.9503
Weight of C_t in effective consumption	λ	0.6
Steady-state G_t to output ratio	\bar{G}/\bar{Y}	0.21
Steady-state labor hours worked	\bar{n}	1/3
Constant TFP	A	1

C Variables and data source

Table 11 Data

Variable	Source
Household final consumption expenditure (constant 2010 US\$)	World Bank national accounts data, and OECD National Accounts data files
Household final consumption expenditure (current US\$)	World Bank national accounts data, and OECD national accounts data files
General government final consumption expenditure (constant 2010 US\$)	World Bank national accounts data, and OECD National accounts data files
General government final consumption expenditure (current US\$)	World Bank national accounts data, and OECD National accounts data files

C.1 Long definition of key variables and sources

- Household final consumption expenditure (formerly private consumption) is the market value of all goods and services, including durable products (such as cars, washing machines, and home computers), purchased by households. It excludes purchases of dwellings but includes imputed rent for owner-occupied dwellings. It also includes payments and fees to governments to obtain permits and licenses. Here, household consumption expenditure includes the expenditures of nonprofit institutions serving households, even when reported separately by the country. Data are in constant 2010 U.S. dollars. Data available at: <https://databank.worldbank.org/source/world-development-indicators>
- Household final consumption expenditure (formerly private consumption) is the market value of all goods and services, including durable products (such as cars, washing machines, and home computers), purchased by households. It excludes purchases of dwellings but includes imputed rent for owner-occupied dwellings. It also includes payments and fees to governments to obtain permits and licenses. Here, household consumption expenditure includes the expenditures of nonprofit institutions serving households, even when reported separately by the country. Data are in current U.S. dollars. Data available at: <https://databank.worldbank.org/source/world-development-indicators>
- General government final consumption expenditure (formerly general government consumption) includes all government current expenditures for purchases of goods and services (including compensation of employees). It also includes most expenditures on national defense and security, but excludes government military expenditures that are part of government capital formation. Data are in constant 2010 U.S. dollars. Data available at: <https://databank.worldbank.org/source/world-development-indicators>
- General government final consumption expenditure (formerly general government consumption) includes all government current expenditures for purchases

Table 12 Estimated Values of IES with Consumption Ratio as Regressand

Panel A	FE	CS-DL
$\hat{\theta}$	0.271*** (0.031)	0.389*** (0.055)
Panel B: diagnostics		
\hat{u}_t : CD p -value	0.000	0.338
$\hat{\varepsilon}_t$: $I(1)$ p -value (ADF)	0.005	0.000
Observation	816	748

Standard error in round brackets. *** represents statistical significance at the 1% level. All estimation equations allow for country-specific effects. CS-DL is the cross-sectional distributed lag estimator by Chudik et al. (2013, 2017). Column [1] assumes cross-sectional independence in the data, while column [2] accounts for cross-sectional dependence in the data

of goods and services (including compensation of employees). It also includes most expenditures on national defense and security, but excludes government military expenditures that are part of government capital formation. Data are in current U.S. dollars. Data available at: <https://databank.worldbank.org/source/world-development-indicators>

- The corresponding prices of private consumption and public consumption are computed as the implicit price deflators, which are constructed by dividing the nominal private and government consumption series defined above by their respective constant price series.

D Estimation of IES with first-differenced variables

We present additional results where the estimation of the IES is carried out using the first-differenced version of the variables in Eq. (6), that is, with the stationary version of the variable of interest. Table 12 presents the estimates from the Fixed Effect (FE) and the CS-DL estimator. In line with the results above, we use the consumption ratio as the dependent variable. As can be observed from the Table, the estimated IES is less than 1 and statistically significant (columns 1 and 2). Consequently, the key conclusion that public and private consumption are Edgeworth complements still holds.

References

- Abrigo MR, Love I (2016) Estimation of panel vector autoregression in Stata. *Stata J* 16:778–804
- Achdou Y, Han J, Lasry J-M, Lions P-L, Moll B (2017) Income and wealth distribution in macroeconomics: a continuous-time approach, working paper 23732, National Bureau of Economic Research
- Amano RA, Wirjanto TS (1997) Intratemporal substitution and government spending. *Rev Econ Stat* 79:605–609
- Amano RA, Wirjanto TS (1998) Government expenditures and the permanent-income model. *Rev Econ Dyn* 1:719–730
- Amendola A, Di Serio M, Fragetta M, Melina G (2020) The Euro-area government spending multiplier at the effective lower bound. *Euro Econ Rev*, 103480

- Baltagi BH, Bresson G, Pirotte A (2008) To pool or not to pool? in *Econom Panel Data*, Springer, 517–546
- Baltagi BH, Griffin JM (1997) Pooled estimators vs. their heterogeneous counterparts in the context of dynamic demand for gasoline. *J Econom* 77:303–327
- Baltagi BH, Pesaran MH (2007) Heterogeneity and cross section dependence in panel data models: theory and applications introduction. *J Appl Economet* 22:229–232
- Bandeira G, Pappa E, Sajedi R, Vella E (2018) Fiscal consolidation in a low-inflation environment: pay cuts versus lost jobs. *Int J Cent Bank* 14:7–52
- Banerjee A, Carrion-i Silvestre JL (2017) Testing for panel cointegration using common correlated effects estimators. *J Time Ser Anal* 38:610–636
- Barrell R, Holland D, Hurst I (2013) Fiscal multipliers and prospects for consolidation. *OECD J: Econom Stud* 2012:71–102
- Baxter M, King RG (1993) Fiscal policy in general equilibrium, *Am Econom Rev*, 315–334
- Boehm CE (2019) Government consumption and investment: does the composition of purchases affect the multiplier? *J Monet Econom*
- Bond S, Leblebiciooğlu A, Schiantarelli F (2010) Capital accumulation and growth: a new look at the empirical evidence. *J Appl Economet* 25:1073–1099
- Bouakez H, Rebei N (2007) Why does private consumption rise after a government spending shock? *Can J Econom/Revue Canadienne d'économique* 40:954–979
- Breitung J (2000) The local power of some unit root tests for panel data, in *Advances in econometrics, Vol. 15: Nonstationary Panels, Panel Cointegration, and Dynamic Panels*, ed. by B. H. Baltagi, Amsterdam: JAI Press, 161–178
- Brown A, Wells G (2008) Substitution between public and private consumption in Australian states, in *Australasian meeting of the econometric society (ESAM08)*, 1–18
- Canova F, Pappa E (2011) Fiscal policy, pricing frictions and monetary accommodation. *Econom Policy* 26:555–598
- Christiano L, Eichenbaum M, Rebelo S (2011) When is the government spending multiplier large? *J Polit Econ* 119:78–121
- Chudik A, Mohaddes K, Pesaran MH, Raissi M (2016): Long-run effects in large heterogeneous panel data models with cross-sectionally correlated errors, in *Essays in honor of Aman Ullah*, Emerald Group Publishing Limited
- Chudik A, Mohaddes K, Pesaran MH, Raissi M (2017) Is there a debt-threshold effect on output growth? *Rev Econ Stat* 99:135–150
- Chudik A, Mohaddes K, Pesaran MH, Raissi M et al (2013) Debt, inflation and growth: robust estimation of long-run effects in dynamic panel data models. Faculty of Economics, University of Cambridge, Tech. rep
- Coenen G, Straub R (2005) Does government spending crowd in private consumption? Theory and empirical evidence for the euro area. *Int Finan* 8:435–470
- Davig T, Leeper EM (2011) Monetary-fiscal policy interactions and fiscal stimulus. *Eur Econ Rev* 55:211–227
- Dawood TC, Francois JN (2018) Substitution between private and government consumption in African economies. *Econ Model* 73:129–139
- De Hoyos RE, Sarafidis V (2006) Testing for cross-sectional dependence in panel-data models. *Stand Genomic Sci* 6:482–496
- Dumitrescu E-I, Hurlin C (2012) Testing for Granger non-causality in heterogeneous panels. *Econ Model* 29:1450–1460
- Eberhardt M (2012) Estimating panel time-series models with heterogeneous slopes. *Stand Genomic Sci* 12:61–71
- Eberhardt M, Presbitero AF (2015) Public debt and growth: heterogeneity and non-linearity. *J Int Econ* 97:45–58
- Eberhardt M, Teal F (2013) Structural change and cross-country growth empirics. *World Bank Econom Rev* 27:229–271
- Eberhardt M, Teal F (2020) The magnitude of the task ahead: macro implications of heterogeneous technology. *Rev Income Wealth* 66:334–360
- Evans P, Karras G (1996) Private and government consumption with liquidity constraints. *J Int Money Financ* 15:255–266
- Fève P, Matheron J, Sahuc J-G (2013) A pitfall with estimated DSGE-based government spending multipliers. *Am Econ J Macroecon* 5:141–178

- Florito R, Kollintzas T (2004) Public goods, merit goods, and the relation between private and government consumption. *Eur Econ Rev* 48:1367–1398
- Forni L, Monteforte L, Sessa L (2009) The general equilibrium effects of fiscal policy: estimates for the euro area. *J Public Econ* 93:559–585
- Francois JN, Keinsley A (2019) The long-run relationship between public consumption and output in developing countries: evidence from panel data. *Econ Lett* 174:96–99
- Galí J, López-Salido JD, Vallés J (2007) Understanding the effects of government spending on consumption. *J Eur Econ Assoc* 5:227–270
- Ganelli G, Tervala J (2009) Can government spending increase private consumption? The role of complementarity. *Econom Lett* 103:5–7
- Granger C (1988) Aggregation of time series variables—a survey. Tech. rep, Federal Reserve Bank of Minneapolis
- Hagedorn M, Manovskii I, Mitman K (2019) The Fiscal Multiplier, Working Paper 25571, National Bureau of Economic Research
- Havráněk T (2015) Measuring intertemporal substitution: the importance of method choices and selective reporting. *J Eur Econ Assoc* 13:1180–1204
- Havráněk T, Sokolova A (2020) Do consumers really follow a rule of thumb? Three thousand estimates from 144 studies say probably not. *Rev Econom Dynam* 35:97–122
- Herzer D, Grimm M (2012) Does foreign aid increase private investment? Evidence from panel cointegration. *Appl Econom* 44:2537–2550
- Herzer D, Morrissey O (2013) Foreign aid and domestic output in the long run. *Rev World Econ* 149:723–748
- Ho T-W (2001) The government spending and private consumption: a panel cointegration analysis. *Int Rev Econom Finan* 10:95–108
- Holly S, Pesaran MH, Yamagata T (2010) A spatio-temporal model of house prices in the USA. *J Econom* 158:160–173
- Holtz-Eakin D, Newey W, Rosen HS (1988) Estimating vector autoregressions with panel data. *Econom J Econom Soc*, 1371–1395
- Hsiao C (2007) Panel data analysis—advantages and challenges. *TEST* 16:1–22
- Hur S-K, Mallick S, Park D (2010) Fiscal policy and crowding out in developing Asia, Asian development bank economics working paper
- Ilzetzki E, Mendoza EG, Végh CA (2013) How big (small?) are fiscal multipliers? *J Monet Econ* 60:239–254
- Im KS, Pesaran MH, Shin Y (2003) Testing for unit roots in heterogeneous panels. *J Econom* 115:53–74
- Jalles JT, Karras G (2021) Private and public consumption: substitutes or complements? *Oxford economic papers*
- Jha S, Mallick SK, Park D, Quising PF (2014) Effectiveness of countercyclical fiscal policy: evidence from developing Asia. *J Macroecon* 40:82–98
- Kao C (1999) Spurious regression and residual-based tests for cointegration in panel data. *J Econom* 90:1–44
- Kao C, Chiang MH (2000) On the estimation and inference of a cointegrated regression in panel data, in *Nonstationary Panels, Panel Cointegration, and Dynamic Panels*, JAI Press, 179–222
- Karras G (1994) Government spending and private consumption: some international evidence. *J Money Credit Bank* 26:9–22
- Kilponen J, Pisani M, Schmidt S, Corbo V, Hledik T, Hollmayr J, Hurtado S, Júlio P, Kulikov D, Lemoine M et al (2019) Comparing Fiscal Consolidation Multipliers across Models in Europe. *Int J Cent Bank* 15:285–320
- Kormendi RC (1983) Government debt, government spending, and private sector behavior. *Am Econ Rev* 73:994–1010
- Kwan YK (2009) The direct substitution between government and private consumption in East Asia. *Fiscal Policy Manag East Asia* 16:45
- Laumer S (2020) Government spending and heterogeneous consumption dynamics. *J Econom Dynam Control*, 103868
- Leeper EM, Traum N, Walker TB (2017) Clearing up the fiscal multiplier morass. *Am Econom Rev* 107:2409–54
- Levin A, Lin C-F, Chu C-SJ (2002) Unit root tests in panel data: asymptotic and finite-sample properties. *J Econom* 108:1–24
- Maddala GS, Wu S (1999) A comparative study of unit root tests with panel data and a new simple test. *Oxford Bull Econ Stat* 61:631–652

- Mark NC, Sul D (2003) Cointegration vector estimation by panel DOLS and long-run money demand. *Oxford Bull Econ Stat* 65:655–680
- Mundlak Y, Butzer R, Larson DF (2008) Heterogeneous technology and panel data: the case of the agricultural production function, The World Bank
- Neal T (2014) Panel cointegration analysis with xtpedroni. *Stand Genomic Sci* 14:684–692
- Ng S, Perron P (1997) Estimation and inference in nearly unbalanced nearly cointegrated systems. *J Econom* 79:53–81
- Ogaki M, Ostry JD, Reinhart CM (1996) Saving behavior in low-and middle-income developing countries a comparison. *IMF Staff Pap* 43:38–71
- Pedroni P (1999) Critical values for cointegration tests in heterogeneous panels with multiple regressors. *Oxford Bull Econ Stat* 61:653–670
- Pedroni P (2001a) Fully modified OLS for heterogeneous cointegrated panels, in *Nonstationary panels, panel cointegration, and dynamic panels*, Emerald Group Publishing Limited, 93–130
- Pedroni P (2001) Purchasing power parity tests in cointegrated panels. *Rev Econ Stat* 83:727–731
- Pedroni P (2019) Panel cointegration techniques and open challenges, In: *Panel data econometrics*, Elsevier, 251–287
- Persyn D, Westerlund J (2008) Error-correction-based cointegration tests for panel data. *Stand Genomic Sci* 8:232–241
- Pesaran MH (2003) Aggregation of linear dynamic models: an application to life-cycle consumption models under habit formation. *Econ Model* 20:383–415
- Pesaran MH (2006) Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica* 74:967–1012
- Pesaran MH (2007) A simple panel unit root test in the presence of cross-section dependence. *J Appl Economet* 22:265–312
- Pesaran MH (2015) Testing weak cross-sectional dependence in large panels. *Economet Rev* 34:1089–1117
- Ramey VA (2011) Identifying government spending shocks: it's all in the timing. *Q J Econ* 126:1–50
- Ramey VA, Shapiro MD (1998) Costly capital reallocation and the effects of government spending. *Carn-Roch Conf Ser Public Policy* 48:145–194
- Sarafidis V, Wansbeek T (2012) Cross-sectional dependence in panel data analysis. *Economet Rev* 31:483–531
- Sarafidis V, Yamagata T, Robertson D (2009) A test of cross section dependence for a linear dynamic panel model with regressors. *J Econom* 148:149–161
- Temple J (1999) The new growth evidence. *J Econom Literat* 37:112–156
- Westerlund J (2007) Testing for error correction in panel data. *Oxford Bull Econ Stat* 69:709–748
- Zubairy S (2014) On fiscal multipliers: estimates from a medium scale DSGE model. *Int Econom Rev* 55:169–195

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.