TI 2010-049/2

Tinbergen Institute Discussion Paper

The Stickiness of Aggregate Consumption Growth in OECD Countries: A Panel Data Analysis

Gerdie Everaert[1]

Lorenzo Pozzi[2]

[1] *SHERPPA, Ghent University;*

[2] *Erasmus University Rotterdam, and Tinbergen Institute.*

Tinbergen Institute

The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

The Stickiness of Aggregate Consumption Growth in OECD Countries: A Panel Data Analysis*

Gerdie Everaert[†1] and Lorenzo Pozzi[‡2]

[1] *SHERPPA, Ghent University*

[2] *Department of Economics, Erasmus University Rotterdam & Tinbergen Institute*

May 6, 2010

Abstract

This paper examines the sources of stickiness in aggregate consumption growth. We first derive a dynamic consumption equation which nests recent developments in consumption theory: rule-of-thumb consumption, habit formation, non-separabilities between both private consumption and hours worked and private consumption and government consumption, intertemporal substitution effects and precautionary savings. Next, we estimate this dynamic consumption equation for a panel of 15 OECD countries over the period 1972-2007 taking into account endogeneity issues and error cross-sectional dependence. To this end, we develop a generalised method of moments version of the common correlated effects pooled estimator and demonstrate its small sample behaviour using Monte Carlo simulations. The estimation results support the labour-consumption complementarity hypothesis, the rule-of-thumb consumption model and the notion that precautionary savings matter for the aggregate economy.

JEL Classification: C23, E21

Keywords: Sticky Consumption, Dynamic Panel, Cross-Sectional Dependence

1 Introduction

The permanent income hypothesis implies that aggregate private consumption follows a random walk (Hall, 1978). Empirical studies show that this random walk hypothesis is not supported by the data since aggregate consumption growth is predictable, at least to some extent. More sophisticated theoretical models reconcile this stylized fact by introducing various forms of stickiness in aggregate consumption growth. Relevant forms are habit formation (Campbell, 1998; Carroll et al., 2008), precautionary savings

*We thank Freddy Heylen for helpful suggestions and constructive comments on an earlier version of this paper.

†I acknowledge financial support from the Interuniversity Attraction Poles Program - Belgian Science Policy, contract no. P5/21.

‡Corresponding author: Department of Economics, P.O. Box 1738, 3000 DR Rotterdam, the Netherlands. Email: pozzi@ese.eur.nl. Website: http://people.few.eur.nl/pozzi.

(Banks et al., 2001), intertemporal substitution effects in response to real interest rate changes (Campbell and Mankiw, 1989), non-separabilities in the utility function between private consumption and government consumption (Evans and Karras, 1998) and between private consumption and hours worked (Basu and Kimball, 2002; Kiley, 2007) and rule-of-thumb consumption (Campbell and Mankiw, 1989, 1990, 1991). Empirically, the most robust finding is the positive impact of aggregate disposable income growth on private consumption growth, which is in general obtained from models incorporating rule-of-thumb consumption. This is the so-called excess sensitivity puzzle of aggregate private consumption. For the other types of stickiness the empirical results are rather mixed. One drawback is that these studies typically focus only on a subset of possible forms of stickiness. Moreover, the empirical analysis is usually restricted to a single country (mainly the US). Studies that present international evidence are Campbell and Mankiw (1991) and Carroll et al. (2008). The disadvantage of these studies is that they use a country-by-country approach. As a result, the additional information in the cross-sectional dimension of the data is not fully exploited. Evans and Karras (1998) and Lopez et al. (2000) use panel data methods but they do not tackle all the complications that arise when estimating aggregate consumption growth equations with macroeconomic data. In particular, they disregard cross-sectional dependence that may stem from the presence of unobserved variables that are common to all countries in the panel.

This paper examines the stickiness of aggregate private consumption growth in a panel of OECD countries over the period 1972-2007. The contribution of the paper to the literature is both theoretical and methodological. Theoretically we present a heterogeneous agent model with consumers who belong to two possible types. Consumers who belong to the first type optimize intertemporally. They form habits since their utility also depends on past consumption. They also have a precautionary savings motive since the variance of their individual labour income affects their consumption decisions. They further substitute consumption intertemporally when confronted with real interest rate changes. Finally, their utility is affected by government consumption and also by the number of hours that they work. Consumers who belong to the second type are rule-of-thumb consumers who consume their entire disposable income in each period. After aggregation over a large population of consumers this model provides an expression for aggregate consumption growth that can be estimated using macroeconomic data. The stickiness factors incorporated in the model lead to the dependence of aggregate private consumption growth on its own lag, on the disposable income to consumption ratio and its square, on the real interest rate, on aggregate government consumption growth, on the growth rate in aggregate hours worked and on aggregate disposable income growth and its lag. These stickiness factors constitute deviations from perfect consumption smoothing as implied by Hall's (1978) random walk hypothesis. Our specification for aggregate consumption growth encompasses most of the recent developments in consumption theory.

And while our specification nests a number of specifications that have been estimated in the literature previously, to the best of our knowledge, no study has yet estimated a specification as general as ours.

Methodologically we estimate the dynamic consumption equation derived in our theoretical model for a panel of 15 OECD countries over the period 1972-2007, making full use of the panel structure of the data. First, we exploit the cross-sectional dimension by pooling information over cross-sections. Second, we exploit the cross-sectional dependence in the data. Standard panel estimators account for unobserved, time-invariant heterogeneity by including individual effects. Recently, the panel literature has emphasized unobserved, time-varying heterogeneity that may stem from omitted common variables that have differential impacts on individual units (see e.g. Coakley et al., 2002; Phillips and Sul, 2003). These latent common variables induce error cross-section dependence and may lead to inconsistent estimates if they are correlated with the explanatory variables. Especially when studying macroeconomic data, such unobserved global variables or shocks (e.g. an international business cycle, oil price shocks, ...) are likely to be the rule rather than the exception (see e.g. Coakley et al., 2006; Westerlund, 2008). Rather than treating the cross-section correlation as a nuisance, we exploit it to correct for a potential omitted variables bias stemming from unobserved common factors. To this end, we use the common correlated effects pooled (CCEP) estimator suggested by Pesaran (2006). The basic idea behind the CCEP estimator is to capture the unobserved common factors by including cross-sectional averages of the dependent and the explanatory variables as additional regressors in the model. Next, we suggest a generalised method of moments (GMM) version of the CCEP estimator to account for endogeneity of the explanatory variables. A small-scaled Monte Carlo simulation shows that in a dynamic panel data model with both endogeneity and error cross-sectional dependence, this CCEP-GMM performs reasonably well for the modest sample size $T = 35, N = 15$ that is available for our empirical analysis, especially when compared to alternative estimators.

The estimation results support the labour-consumption complementarity hypothesis, the rule-of-thumb consumption model and the notion that precautionary savings matter for the aggregate economy. We find little or no support for habit formation, non-separabilities between private consumption and government consumption and intertemporal substitution effects. Taking into account endogeneity and cross-sectional dependence proves to be important as it has a marked effect on the coefficient estimates.

The paper is structured as follows. In section 2 we derive a dynamic equation for aggregate private consumption growth from a model that encompasses most of the recent developments in the consumption literature. In section 3 we review the different estimators that can be used to estimate this equation in a panel of OECD countries and investigate their small sample properties in a Monte Carlo experiment. Section 4 presents the results from the estimation of the consumption growth equation with the different panel data estimators. Section 5 concludes.

2 The model

Consider a heterogeneous agent economy with two consumer types (see Campbell and Mankiw, 1989, 1990, 1991). Type 1 are intertemporally optimizing consumers, described in section 2.2, while type 2 are rule-of-thumb consumers, described in section 2.3. Consumers within each type are heterogeneous in the sense that they experience different shocks to labour income and that they work different amounts of hours. Consumption in the model is driven by three exogenous but potentially correlated stochastic processes: individual labour income, individual hours worked, and aggregate government consumption. Only the process for individual labour income is explicitly specified in section 2.1. Consumption is also driven by the exogenous economy-wide interest rate which is time-varying but not stochastic. In the remainder, upper case variables are aggregate economy-wide while lower case variables are individual consumer-specific, i.e. they take on different values for each consumer both within and across consumer types.

2.1 Individual labour income

Individual log labour income of each consumer in the economy (belonging to either type 1 or type 2) is exogenous and is assumed to follow a random walk process

$$y_t = y_{t-1}\varepsilon_t, \tag{1}$$

$$\ln y_t = \ln y_{t-1} + \ln \varepsilon_t, \tag{2}$$

with

$$E_{t-1} \ln \varepsilon_t = 0, \tag{3}$$

$$\text{and} \qquad V_{t-1} \ln \varepsilon_t = \sigma^2_{\ln \varepsilon}. \tag{4}$$

From this we can approximate the conditional mean and variance of the error term ε_t as $E_{t-1}\varepsilon_t = z$ with $z \approx 1 + \frac{1}{2}\sigma^2_{\ln \varepsilon} > 1$ and $V_{t-1}\varepsilon_t \approx \sigma^2_{\ln \varepsilon}$. We refer to Appendix A.1 for the derivation.

2.2 Type 1: intertemporally optimizing consumers

Consumers belonging to type 1 are intertemporally optimizing consumers who accumulate and decumulate wealth. Let w_t denote individual resources in period t, defined as individual assets augmented with interest on these assets plus individual labour income, i.e. $w_t \equiv (1+R_t)a_{t-1}+y_t$ where a_{t-1} are individual assets at the end of period $t-1$ and where R_t is the exogenous economy-wide time-varying but risk-free interest rate (see e.g. Deaton, 1991). The law of motion for w_t is then given by

$$w_t = (1+R_t)(w_{t-1} - c_{t-1}) + y_t, \tag{5}$$

where use is made of the budget constraint which implies that $a_{t-1} = w_{t-1} - c_{t-1}$, with c_{t-1} denoting the consumption level of an individual consumer at time $t-1$. Note from this equation that the anticipated part of resources equals $\overline{w}_t = E_{t-1}w_t = (1+R_t)(w_{t-1} - c_{t-1}) + y_{t-1}z$. This is obtained by noting that $E_{t-1}y_t = y_{t-1}E_{t-1}\varepsilon_t = y_{t-1}z$. The shock to resources then equals $w_t - E_{t-1}w_t = w_t - \overline{w}_t = y_t - y_{t-1}z = y_{t-1}(\varepsilon_t - z) = \eta_t$ with $E_{t-1}\eta_t = 0$ and $V_{t-1}\eta_t = y_{t-1}^2 V_{t-1}(\varepsilon_t - z) = y_{t-1}^2 V_{t-1}\varepsilon_t = y_{t-1}^2 \sigma_{\ln \varepsilon}^2$.

The contemporaneous utility function of an individual consumer who belongs to type 1 and who decides on consumption c_t is given by

$$u(c_t) = \frac{1}{1-\theta}\left[\frac{c_t}{c_{t-1}^{\beta}}\right]^{1-\theta} e^{\gamma \ln h_t} e^{\pi \ln G_t}, \tag{6}$$

where h_t is the exogenous stochastic number of hours worked of this individual, G_t denotes exogenous stochastic aggregate government consumption, $\theta > 0$ is the coefficient of relative risk aversion, $\beta \geq 0$ is the habit parameter (Campbell, 1998), and γ and π capture respectively the impact of hours worked (Campbell and Mankiw, 1990) and government consumption (Evans and Karras, 1998) on the marginal utility of private consumption. This is a utility function of the King-Plosser-Rebelo type as used e.g. by Basu and Kimball (2002, p.5). When $\gamma > 0$ (< 0) hours worked and private consumption are complements (substitutes). When $\pi > 0$ (< 0) government consumption and private consumption are complements (substitutes). Note that $\gamma > 0$ and $\pi < 0$ does not imply that hours worked increase and government consumption decrease total utility of consumption since a function $\phi(h_t, G_t)$ could be added to the utility function (with $\phi_h < 0$ and $\phi_G > 0$) without changing the first-order condition.

The first-order condition of an individual consumer belonging to type 1 is given by

$$u'(c_{t-1}) = \left(\frac{1+R_t}{1+\delta}\right) E_{t-1}u'(c_t),$$

where $0 < \delta < 1$ is the rate of time preference. Following Banks et al. (2001) we assume that optimal consumption c_t is approximately proportional to resources w_t, i.e. $c_t \approx \alpha w_t$. Substituting eq.(6) into the first-order condition gives

$$c_{t-1}^{-\theta}\left(\frac{c_{t-1}}{c_{t-2}}\right)^{-\beta(\theta-1)}e^{\gamma\ln h_{t-1}}e^{\pi\ln G_{t-1}}=\left(\frac{1+R_t}{1+\delta}\right)E_{t-1}\left(f(w_t,\ln h_t,\ln G_t)\right), \tag{7}$$

where $f(w_t,\ln h_t,\ln G_t)=(\alpha w_t)^{-\theta}e^{\gamma\ln h_t}e^{\pi\ln G_t}$.

We derive an expression for $f(w_t,\ln h_t,\ln G_t)$ in the proximity of certainty, i.e. in the situation where the shocks to the exogenous stochastic driving processes of the model (individual labour income, individual hours worked, aggregate government consumption) equal zero. In particular, as in Banks et al. (2001), we take a second-order Taylor approximation of $f(w_t,\ln h_t,\ln G_t)=(\alpha w_t)^{-\theta}e^{\gamma\ln h_t}e^{\pi\ln G_t}$ around $c_t=\alpha w_t=\alpha\overline{w}_t$ (where the last step follows from setting $\eta_t=0$), around $\ln h_t=\overline{\ln h}_t$ (where $\overline{\ln h}_t=E_{t-1}\ln h_t$), and around $\ln G_t=\overline{\ln G}_t$ (where $\overline{\ln G}_t=E_{t-1}\ln G_t$). We refer to Appendix A.2 for the derivation. After taking expectations at time $t-1$ of the approximation for $f(w_t,\ln h_t,\ln G_t)$ we obtain

$$E_{t-1}\left(f(w_t,\ln h_t,\ln G_t)\right)\approx\left[(\alpha\overline{w}_t)^{-\theta}e^{\gamma\overline{\ln h}_t}e^{\pi\overline{\ln G}_t}\right]p_{t-1}, \tag{8}$$

where

$$\begin{aligned} p_{t-1} &= 1+\frac{1}{2}\alpha^2\theta(1+\theta)\left(\frac{y_{t-1}}{\alpha\overline{w}_t}\right)^2\sigma^2_{\ln\varepsilon}+\frac{1}{2}\gamma^2\sigma^2_{\ln h}+\frac{1}{2}\pi^2\sigma^2_{\ln G} \\ &\quad -\alpha\theta\gamma\left(\frac{y_{t-1}}{\alpha\overline{w}_t}\right)\sigma_{\varepsilon\ln h}-\alpha\theta\pi\left(\frac{y_{t-1}}{\alpha\overline{w}_t}\right)\sigma_{\varepsilon\ln G}+\gamma\pi\sigma_{\ln h\ln G}, \end{aligned} \tag{9}$$

with $\sigma^2_{\ln\varepsilon}$ as defined in eq.(4), with $\sigma^2_{\ln h}=E_{t-1}(\ln h_t-\overline{\ln h}_t)^2$, with $\sigma^2_{\ln G}=E_{t-1}(\ln G_t-\overline{\ln G}_t)^2$, with $\sigma_{\varepsilon\ln h}=E_{t-1}\left[(\varepsilon_t-z)(\ln h_t-\overline{\ln h}_t)\right]$, with $\sigma_{\varepsilon\ln G}=E_{t-1}\left[(\varepsilon_t-z)(\ln G_t-\overline{\ln G}_t)\right]$, and with $\sigma_{\ln h\ln G}=E_{t-1}\left[(\ln h_t-\overline{\ln h}_t)(\ln G_t-\overline{\ln G}_t)\right]$.

Substituting eq.(8) into eq.(7) and solving for c_{t-1} gives

$$c_{t-1}=\left(\frac{1+R_t}{1+\delta}\right)^{-\frac{1}{\theta}}(\alpha\overline{w}_t)e^{-\frac{\gamma}{\theta}E_{t-1}\Delta\ln h_t}e^{-\frac{\pi}{\theta}E_{t-1}\Delta\ln G_t}\left(\frac{c_{t-1}}{c_{t-2}}\right)^{\frac{\beta(1-\theta)}{\theta}}p_{t-1}^{-\frac{1}{\theta}}. \tag{10}$$

Now we divide $c_t=\alpha w_t$ by c_{t-1} as given in eq.(10) to obtain

$$\frac{c_t}{c_{t-1}}=\left(1+\frac{\eta_t}{\overline{w}_t}\right)\left(\frac{1+R_t}{1+\delta}\right)^{\frac{1}{\theta}}e^{\frac{\gamma}{\theta}E_{t-1}\Delta\ln h_t}e^{\frac{\pi}{\theta}E_{t-1}\Delta\ln G_t}\left(\frac{c_{t-1}}{c_{t-2}}\right)^{\frac{\beta(\theta-1)}{\theta}}p_{t-1}^{\frac{1}{\theta}}, \tag{11}$$

where we use the result that $\frac{w_t}{\overline{w}_t}=1+\frac{\eta_t}{\overline{w}_t}$.

Taking logs of both sides of eq.(11) we obtain

$$\Delta\ln c_t=-\frac{\delta}{\theta}+\frac{\beta(\theta-1)}{\theta}\Delta\ln c_{t-1}+\frac{\gamma}{\theta}\Delta\ln h_t+\frac{\pi}{\theta}\Delta\ln G_t+\frac{1}{\theta}R_t+\frac{1}{\theta}\ln p_{t-1}+\psi_t, \tag{12}$$

where we use the approximation $\ln(1+x) \approx x$ and where $\psi_t = \frac{\eta_t}{\overline{w}_t} - \frac{\gamma}{\theta}\left(\Delta \ln h_t - E_{t-1}\Delta \ln h_t\right) - \frac{\pi}{\theta}\left(\Delta \ln G_t - E_{t-1}\Delta \ln G_t\right)$. Note that $E_{t-1}\psi_t = 0$.

Note that from eq.(9) we can write

$$\ln p_{t-1} = k_0 + k_1 \left(\frac{y_{t-1}}{c_{t-1}}\right) + k_2 \left(\frac{y_{t-1}}{c_{t-1}}\right)^2, \tag{13}$$

where again we use the approximation $\ln(1+x) \approx x$, where we replace $\alpha\overline{w}_t$ by c_{t-1} (see Banks et al., 2001), where $k_0 = \frac{1}{2}\gamma^2\sigma^2_{\ln h} + \frac{1}{2}\pi^2\sigma^2_{\ln G} + \gamma\pi\sigma_{\ln h \ln G}$, where $k_1 = -\alpha\theta(\gamma\sigma_{\varepsilon \ln h} + \pi\sigma_{\varepsilon \ln G})$ and where $k_2 = \frac{1}{2}\alpha^2\theta\left(1+\theta\right)\sigma^2_{\ln \varepsilon}$.

Substituting eq.(13) into eq.(12) gives the growth rate of consumption of an individual consumer belonging to type 1,

$$\begin{aligned} \Delta \ln c_t \quad = \quad & \left(\frac{k_0 - \delta}{\theta}\right) + \frac{\beta(\theta-1)}{\theta}\Delta \ln c_{t-1} + \frac{\gamma}{\theta}\Delta \ln h_t + \frac{\pi}{\theta}\Delta \ln G_t \\ & + \frac{1}{\theta}R_t + \frac{k_1}{\theta}\left(\frac{y_{t-1}}{c_{t-1}}\right) + \frac{k_2}{\theta}\left(\frac{y_{t-1}}{c_{t-1}}\right)^2 + \psi_t. \end{aligned} \tag{14}$$

2.3 Type 2: rule-of-thumb consumers

Consumers belonging to type 2 are rule-of-thumb consumers who consume their entire disposable labour income in each period due to for instance myopia (see Flavin, 1985) or liquidity constraints (see Jappelli and Pagano, 1989; Campbell and Mankiw, 1990). Hence they do not accumulate or decumulate wealth. For these consumers $c_t = y_t$ for all t. Hence, the growth rate of consumption of an individual consumer belonging to type 2 is given by,

$$\Delta \ln c_t = \Delta \ln y_t. \tag{15}$$

2.4 Aggregate consumption growth

The growth rate of total consumption of type 1 consumers is calculated by summing eq.(14) over all type 1 consumers. We obtain,

$$\Delta \ln C_{1,t} = \left[\begin{array}{c} \left(\frac{k_0-\delta}{\theta}\right) + \frac{\beta(\theta-1)}{\theta}\Delta \ln C_{1,t-1} + \frac{\gamma}{\theta}\Delta \ln H_{1,t} + \frac{\pi}{\theta}\Delta \ln G_t \\ +\frac{1}{\theta}R_t + \frac{k_1}{\theta}\left(\frac{Y_{1,t-1}}{C_{1,t-1}}\right) + \frac{k_2}{\theta}\left(\frac{Y_{1,t-1}}{C_{1,t-1}}\right)^2 + \Psi_t \end{array} \right], \tag{16}$$

where $C_{1,t}$ denotes total consumption of type 1 consumers[1], $H_{1,t}$ denotes total hours worked by type 1 consumers, $Y_{1,t}$ denotes total labour income earned by type 1 consumers, $\Psi_t = \int \psi_t d\Gamma_t$ is the aggregate innovation where Γ_t is the measure of consumers over their state variables (see Gourinchas and Parker,

[1]Note that $C_{1,t}$ is a sum not an average.

2001). For the aggregate innovation we have $E_{t-1}\Psi_t = 0$.[2]

The growth rate of total consumption of type 2 consumers is calculated by summing eq.(15) over all type 2 consumers. We obtain,

$$\Delta \ln C_{2,t} = \Delta \ln Y_{2,t}, \tag{17}$$

where $C_{2,t}$ denotes total consumption of type 2 consumers[3], and $Y_{2,t}$ denotes total labour income earned by type 2 consumers.

Following Campbell and Mankiw (1990) in Appendix A.3 we replace the growth rate of labour income of type 2 consumers $\Delta \ln Y_{2,t}$ by $(1-\lambda)\Delta \ln Y_t$ where λ is the fraction of labour income in the economy accruing to type 1 consumers with $0 < \lambda \leq 1$. The growth rate of total consumption C_t in the economy is then given by

$$\Delta \ln(C_t) = \Delta \ln(C_{1,t}) + (1-\lambda)\Delta \ln(Y_t). \tag{18}$$

From this note that

$$\Delta \ln(C_{1,t}) = \Delta \ln(C_t) - (1-\lambda)\Delta \ln(Y_t), \tag{19}$$

and that

$$\Delta \ln(C_{1,t-1}) = \Delta \ln(C_{t-1}) - (1-\lambda)\Delta \ln(Y_{t-1}). \tag{20}$$

In Appendix A.4 we calculate approximations for $(Y_{1,t-1}/C_{1,t-1})$ and $(Y_{1,t-1}/C_{1,t-1})^2$, namely

$$(Y_{1,t-1}/C_{1,t-1}) = 1 - \lambda^{-1} + \lambda^{-1}(Y_{t-1}/C_{t-1}), \tag{21}$$

and

$$(Y_{1,t-1}/C_{1,t-1})^2 = 1 - \lambda^{-1} + \lambda^{-1}\left(Y_{t-1}/C_{t-1}\right)^2. \tag{22}$$

We then assume that the growth rate in total hours worked for consumers of type 1 is proportional to the growth rate of hours worked in the entire economy where the factor of proportionality equals λ. Hence hours worked are allocated to both consumer types in the same proportion as labour income. Hence,

$$\Delta \ln H_{1,t} = \lambda \Delta \ln H_t \tag{23}$$

where H_t denotes total hours worked in the economy.

By substituting eqs.(19),(20),(21), (22), and (23) into (16) , we obtain an expression for the growth

[2]Note that in the aggregation we assume that $\int \frac{y_{t-1}}{c_{t-1}} d\Gamma_{t-1} \approx \left(\frac{Y_{1,t-1}}{C_{1,t-1}}\right)$ and $\int \left(\frac{y_{t-1}}{c_{t-1}}\right)^2 d\Gamma_{t-1} \approx \left(\frac{Y_{1,t-1}}{C_{1,t-1}}\right)^2$.

[3]Note that $C_{2,t}$ is a sum not an average.

rate in total consumption

$$\begin{aligned}\Delta \ln C_t &= a_0 + a_1 \Delta \ln C_{t-1} + a_2 \Delta \ln H_t + a_3 \Delta \ln G_t + a_4 R_t \\ &\quad + a_5 \left(\frac{Y_{t-1}}{C_{t-1}}\right) + a_6 \left(\frac{Y_{t-1}}{C_{t-1}}\right)^2 + a_7 \Delta \ln Y_t + a_8 \Delta \ln Y_{t-1} + \Psi_t,\end{aligned} \tag{24}$$

where $a_0 = \left(\frac{k_0-\delta}{\theta}\right) + \frac{k_1}{\theta}(1-\lambda^{-1}) + \frac{k_2}{\theta}(1-\lambda^{-1})$, $a_1 = \frac{\beta(\theta-1)}{\theta}$, $a_2 = \lambda\left(\frac{\gamma}{\theta}\right)$, $a_3 = \frac{\pi}{\theta}$, $a_4 = \frac{1}{\theta}$, $a_5 = \lambda^{-1}\left(\frac{k_1}{\theta}\right)$, $a_6 = \lambda^{-1}\left(\frac{k_2}{\theta}\right)$, $a_7 = (1-\lambda)$, $a_8 = -(1-\lambda)\left(\frac{\beta(\theta-1)}{\theta}\right)$ and where $E_{t-1}\Psi_t = 0$.

It is straightforward to show that when $\lambda = 1$ (i.e. when all labour income is earned by type 1 optimizing consumers) then eq.(24) collapses to eq.(16). In Appendix A.5 we further show that for values of λ approaching 0 (i.e. when all labour income is earned by type 2 rule-of-thumb consumers) eq.(24) collapses to $\Delta \ln C_t = \Delta \ln Y_t$, i.e. $\lim_{\lambda \to 0} \Delta \ln C_t = \Delta \ln Y_t$.

Our consumption equation (24) nests most of the recent developments in consumption theory. The parameter $a_1 \geq 0$ reflects habit formation. It is determined by the structural parameter capturing habits, i.e. $\beta \geq 0$. Non-zero values for a_2 and a_3 reflect non-separabilities between private consumption and hours worked and private consumption and government consumption respectively. These parameters are determined by the structural parameters γ respectively π. The parameter $a_4 > 0$ reflects intertemporal substitution effects in consumption caused by interest rate changes. It is determined by the structural parameter $1/\theta$, i.e. the inverse of the coefficient of relative risk aversion $\theta > 0$. The parameter $a_6 > 0$ reflects the impact of precautionary savings on aggregate consumption growth. It depends on the structural parameters θ (the coefficient of relative risk aversion which also reflects the degree of prudence) and $\sigma^2_{\ln\varepsilon}$ (the variance of shocks to individual labour income).[4] The parameter a_7 ($0 \leq a_7 < 1$) reflects rule-of-thumb consumption (liquidity constraints, myopia). It depends on the structural parameter λ (where $0 < \lambda \leq 1$). The parameter a_8 ($a_8 \leq 0$) reflects the need for an additional term $\Delta \ln Y_{t-1}$ in the equation to correct the for the fact that only consumers belonging to type 1 form habits. Finally, the coefficient a_5 depends on the non-separabilities between private consumption and hours worked and private consumption and government consumption, on the correlation between the labour income shocks and shocks to hours worked, and on the correlation between the labour income shocks and shocks in

[4]The importance of this variance for aggregate consumption growth shows why the use of a representative agent model would be inappropriate if we want to take the possibility of a precautionary savings effect at the aggregate level seriously. In a representative agent model the parameter $\sigma^2_{\ln\varepsilon}$ would be the variance of aggregate labour income shocks. This variance is very small so that in a representative agent model we would a priori impose $a_6 \approx 0$, i.e. the variance would be too small to explain aggregate consumption growth (see Gourinchas and Parker, 2001). To see this note that the impact of precautionary savings on aggregate consumption growth in eq.(24) is captured by $\lambda^{-1}\frac{k_2}{\theta} = \lambda^{-1}\frac{1}{2}\alpha^2(1+\theta)\sigma^2_{\ln\varepsilon}$. If all consumers are of type 1 ($\lambda = 1$) then under plausible values for risk aversion ($0.5 < \theta < 5$) precautionary savings motives are irrelevant if $\sigma^2_{\ln\varepsilon}$ denotes the variance of shocks to aggregate labour income (e.g. in annual US data we find that this variance is a mere 0.0001). If, as is the case in our model, $\sigma^2_{\ln\varepsilon}$ denotes the variance of shocks to consumer-specific labour income then it will be much larger (e.g. about 100 times larger in annual US data) and the impact of precautionary savings on aggregate consumption growth will be non-negligible.

aggregate government consumption. It has little intuitive appeal. Note that some of the coefficients in eq.(24) could be given other interpretations. A positive coefficient a_1 on lagged aggregate consumption growth could also be the result of the presence of consumers who are inattentive to macro developments (see Carroll et al., 2008). Further, a positive coefficient a_7 on current aggregate labour income growth could also be the result of consumers who are imperfectly informed about the aggregate economy (see Goodfriend, 1992; Pischke, 1995).[5]

To the best of our knowledge no study has yet estimated a specification as general as ours. The equation nests however a number of specifications that have been estimated in the literature previously. Campbell and Mankiw (1990) conduct regressions on a version of eq.(24) with restrictions $a_1 = a_5 = a_6 = a_8 = 0$ (with $\Delta \ln Y$ always included and either $\Delta \ln H$, $\Delta \ln G$, or R added as an additional regressor). Evans and Karras (1998) estimate a version of eq.(24) with restrictions $a_1 = a_2 = a_4 = a_5 = a_6 = a_8 = 0$ (with $\Delta \ln Y$ and $\Delta \ln G$ included). Basu and Kimball (2002) estimate a version of eq.(24) with restrictions $a_1 = a_3 = a_5 = a_6 = a_8 = 0$ (with $\Delta \ln H$, $\Delta \ln Y$, and R included). Kiley (2007) estimates a version of eq.(24) with restrictions $a_3 = a_5 = a_6 = a_8 = 0$ (with $\Delta \ln H$, $\Delta \ln Y$, $\Delta \ln C_{t-1}$, and R included). Carroll et al. (2008) estimate a version of eq.(24) with restrictions $a_2 = a_3 = a_4 = a_5 = a_6 = a_8 = 0$ (with $\Delta \ln C_{t-1}$ and $\Delta \ln Y$ included). They also add the lagged aggregate wealth level of households in an ad hoc way to their equation. This serves as a proxy for both intertemporal substitution effects and precautionary savings effects which in our equation are captured through the variables R and $\left(\frac{Y}{C}\right)^2$ respectively and which we obtain as an outcome of optimization.[6]

The error term μ_t is assumed to be unpredictable based on lagged information. Three features that are not incorporated in the model could lead to a violation of this assumption and to the occurrence of autocorrelation of the moving average form in the error term μ_t. First, Campbell and Mankiw (1990) note that transitory consumption and measurement error could lead to an $MA(1)$ structure of the error term, i.e. a negative $MA(1)$ coefficient. Second, Working (1960) shows that if consumption decisions are more frequent than observed data then an $MA(1)$ component could be present in consumption growth, i.e. a positive $MA(1)$ coefficient.[7] Third, if durable consumption components are present in C_t this induces negative autocorrelation in $\Delta \ln C_t$ since durable consumption growth is slightly negatively autocorrelated (see Mankiw, 1982). This negative autocorrelation could be reflected in less positive values for a_1 or in

[5] On the basis of macro data alone - on which the empirical analysis of this paper is based - it is not possible to distinguish the interpretations derived from the model from these alternative possibilities.

[6] With respect to precautionary savings note that the factor $\left(\frac{Y}{C}\right)^2$ implies that it is the variance of income shocks $\sigma^2_{\ln \varepsilon}$ *relative* to resources (here proxied by C) that determines the strength of the precautionary savings motive. The lower resources (and therefore consumption) compared to income, the larger the precautionary savings effect, i.e. precaution matters more if wealth is low.

[7] With habit formation this amounts to an $MA(2)$ process but as noted by Carroll et al. (2008) the $MA(2)$ coefficient is generally small.

a more negative $MA(1)$ coefficient in the error term.

3 Econometric methodology

The model for aggregate consumption growth given in eq.(24) can be written in the form of a first-order autoregressive panel data model

$$y_{it} = \alpha_i + \rho y_{i,t-1} + \beta' x_{it} + \mu_{it}, \qquad i = 1, 2, \ldots, N, \quad t = 2, \ldots T, \tag{25}$$

where $y_{it} = \Delta \ln C_{it}$, $x_{it} = (x'_{1it}, x'_{2it})'$ is a vector of explanatory variables with $x_{1it} = \left(\left(\frac{Y_{i,t-1}}{C_{i,t-1}}\right), \left(\frac{Y_{i,t-1}}{C_{i,t-1}}\right)^2, \Delta \ln Y_{i,t-1}\right)'$ and $x_{2it} = (\Delta \ln H_{it}, \Delta \ln G_{it}, R_{it}, \Delta \ln Y_{it})'$. Unobserved time-invariant heterogeneity is captured by the individual effect α_i.

The cases of cross-sectional independence and cross-sectional dependence in the error terms μ_{it} are considered separately in the sections 3.1 and 3.2. Given that the WG transformation suggested in section 3.1 requires $T \to \infty$ and the CCEP estimator suggested in section 3.2 requires $N, T \to \infty$ jointly, section 3.3 reports the results of a small Monte Carlo experiment which evaluates the performance of the suggested estimators for the limited sample of $T = 35$ and $N = 15$ that is available to us for the empirical analysis in section 4.

3.1 Model with cross-sectional independence

As noted in section 2.4, there are various reasons that could lead to the occurrence of autocorrelation of the MA form in the error term of eq.(24). Therefore, we allow μ_{it} in the empirical model in eq.(25) to follow an $MA(q)$ process

$$\mu_{it} = \phi(L)\, \varepsilon_{it}, \tag{26}$$

where $\phi(L) = 1 + \phi_1 L + \ldots + \phi_q L^q$ is a lag polynomial of order q and ε_{it} is an idiosyncratic error term satisfying the following error condition.

Assumption A1 *(Error condition)*

(a) $E(\varepsilon_{it}) = 0$ *for all* i *and* t*;*

(b) $E(\varepsilon_{it}\varepsilon_{js}) = 0$ *for either* $i \neq j$*, or* $t \neq s$*, or both;*

(c) $E(\varepsilon_{it}\alpha_j) = 0$ *for all* i*,* j *and* t*.*

A1(a) and **A1**(b) state that ε_{it} is a mean zero error process which is mutually uncorrelated over time and over cross sections. **A1**(c) states that the individual effects are exogenous.

With respect to the explanatory variables we make the following assumptions.

Assumption A2 *(Explanatory variables)*

(a) $E(x_{1jt}\varepsilon_{is}) = 0$ *for all* i, j, t *and* $s \geq t$*;*

(b) $E(x_{2jt}\varepsilon_{is}) = 0$ *for all* i, j, t *and* $s > t$*;*

(c) $E(\alpha_i x_{it}) =$ *unknown for all* i *and* t*.*

A2(a) states that the variables in x_{1it} are predetermined while **A2**(b) allows the variables in x_{2it} to be endogenous. **A2**(c) allows the explanatory variables to be correlated with α_i.

In the remainder of this section we outline our approach to estimate the panel model in equations (25)-(26) under **A1** and **A2**. We start by arguing that, given the moderately large T dimension of our panel dataset (see section 4), the WG transformation is valid to eliminate the individual effects α_i. Next, we set up a GMM version of the standard WG estimators to take into account the possible $MA(q)$ structure of μ_{it} and the endogeneity of x_{2it}.

3.1.1 WG estimator

The standard within-groups (WG) estimator eliminates the individual effects α_i by transforming the model in (25) into deviations from individual means

$$\widetilde{y}_{it} = \rho\widetilde{y}_{i,t-1} + \beta'\widetilde{x}_{it} + \widetilde{\mu}_{it}, \tag{27}$$

where $\widetilde{y}_{it} = y_{it} - \overline{y}_i$ with $\overline{y}_i = T^{-1}\sum_{t=1}^{T} y_{it}$ and similarly for the other variables.

Even when abstracting from the $MA(q)$ structure of μ_{it} and from the endogeneity of x_{2it}, the least squares (LS) estimation of equation (27) yields inconsistent estimates for $N \rightarrow \infty$ as the within transformed lagged dependent variable $\widetilde{y}_{i,t-1}$ is correlated with the within transformed error term $\widetilde{\mu}_{it}$. However, this dynamic panel data bias disappears for $T \rightarrow \infty$ (see e.g. Nickell, 1981) as $\underset{T\rightarrow\infty}{\text{plim}}\, \overline{\mu}_{it} = 0$ such that $\widetilde{\mu}_{it} \xrightarrow{p} \mu_{it}$. More specifically, the Monte Carlo experiments in Judson and Owen (1999) show that the bias sharply decays when the time horizon exceeds 20 periods, with the bias being negligible when T is 30. Given that our dataset covers the period 1972-2007, we therefore use the WG instead of the first-difference transformation to eliminate the individual effects as first-difference and even system GMM estimators have a relatively large standard error (see e.g. Arellano and Bond, 1991; Kiviet, 1995) and may suffer from a serious weak instruments problem (see e.g. Blundell and Bond, 1998; Bun and Windmeijer, 2007).

3.1.2 WG-GMM estimator

Even as $T \to \infty$, the WG estimator is inconsistent, though, as **A2** states that the variables in x_{2it} are endogenous while the $MA(q)$ structure of μ_{it} in equation (26) implies that the predetermined variables in x_{1it} are also correlated with μ_{it}. Therefore, we estimate (27) using IV methods. More specifically, provided T is sufficiently large, such that $\widetilde{\mu}_{it} \xrightarrow{p} \mu_{it}$, valid WG orthogonality conditions[8] are

$$E\left(\widetilde{y}_{i,t-l}\widetilde{\mu}_{it}\right) = 0 \quad \text{for each} \quad t = l+1,\dots,T \text{ and } l \geq q+1, \tag{28a}$$

$$E\left(\widetilde{x}_{1i,t-l}\widetilde{\mu}_{it}\right) = 0 \quad \text{for each} \quad t = l+1,\dots,T \text{ and } l \geq q, \tag{28b}$$

$$E\left(\widetilde{x}_{2i,t-l}\widetilde{\mu}_{it}\right) = 0 \quad \text{for each} \quad t = l+1,\dots,T \text{ and } l \geq q+1. \tag{28c}$$

Note that the number of moment conditions suggested in (28a)-(28c) increases dramatically in T as they are valid for each t and l. For the moderately large T panel we have in mind, this results in a problem of too many instruments. This instrument proliferation may cause small sample problems like (i) coefficient estimates that are biased towards those of non-instrumenting estimators because of overfitting of the endogenous variables, (ii) imprecise estimates of the optimal weighting matrix, (iii) a downward bias of the two-step standard errors and (iv) a weak Hansen test on instrument validity (see among others Tauchen, 1986; Ziliak, 1997; Roodman, 2009). Therefore we reduce the instrument count using two techniques described in Roodman (2009) and applied by, among others, Pozzi and Malengier (2007). First, we do not exploit all the available linear orthogonality conditions but truncate the set of available instruments at the first L available lags rather than using all available lags. This results in the following reduced set of moment conditions

$$E\left(\widetilde{y}_{i,t-l}\widetilde{\mu}_{it}\right) = 0 \quad \text{for each} \quad q+1 \leq l \leq L+q, \tag{29a}$$

$$E\left(\widetilde{x}_{1i,t-l}\widetilde{\mu}_{it}\right) = 0 \quad \text{for each} \quad q \leq l \leq L+q-1, \tag{29b}$$

$$E\left(\widetilde{x}_{2i,t-l}\widetilde{\mu}_{it}\right) = 0 \quad \text{for each} \quad q+1 \leq l \leq L+q. \tag{29c}$$

Second, we further limit the number of moment conditions by taking linear combinations. More specifically, we combine all moment conditions for a given instrument lag distance by summing them over T. This new set of moment conditions embodies the same belief about orthogonality of the instruments and the errors but differs in that we ask the GMM estimator to minimize the magnitude of the empirical moments $\sum_i \sum_t y_{i,t-l}\mu_{it}$ for each l rather that $\sum_i y_{i,t-l}\mu_{it}$ for each t and l separately. As such, each lag distance produces a single instrument for each variable. As in Holtz-Eakin et al. (1988) we avoid the

[8]See for instance Meghir and Pistaferri (2004) for an application of the WG-GMM estimator

trade-off between instrument lag depth and sample depth by zeroing out missing observations of lags.

3.2 Model with cross-sectional dependence

Following the recent panel literature, we extend the error process in equation (26) to allow for a multi-factor structure

$$\mu_{it} = \gamma_i' f_t + \phi(L)\,\varepsilon_{it}, \tag{30}$$

in which f_t is an $m \times 1$ vector of unobserved common variables and ε_{it} satisfies **A1**. This error structure is quite general as it allows for an unknown (but fixed) number of unobserved common components with heterogeneous factor loadings (heterogeneous cross-sectional dependence). As such, it nests common time effects or time dummies (homogeneous cross-sectional dependence) as a special case.

Assumption A3 *(Cross-sectional dependence)*

(a) The unobserved factors f_t can follow general covariance stationary processes;

(b) $E(f_t \varepsilon_{is}) = 0$ for all i, t and s;

(c) $E(f_t x_{is}) =$ unknown for all i, t and s.

A3 states that the unobserved factors in f_t are exogenous but is quite general as it allows f_t to exhibit rich dynamics and to be correlated with x_{it}. As **A1** states that ε_{it} is uncorrelated over cross sections, any dependence across countries is restricted to the common factors.

The most obvious implication of ignoring error cross-sectional dependence is that it increases the variation of standard panel data estimators. Phillips and Sul (2003) for instance show that if there is high cross-sectional correlation there may not be much to gain from pooling the data. However, cross-sectional dependence can also introduce a bias and even result in inconsistent estimates. For a static panel data model, the Monte Carlo simulations in Pesaran (2006) reveal that the WG estimator ignoring the error component structure proposed in (30) is seriously biased. Essentially, as the unobserved factors are allowed to be correlated with the explanatory variables (see **A3**), this is an omitted variables bias which does not disappear as $T \to \infty$, $N \to \infty$ or both. So the naive WG estimator is biased and even inconsistent in this case. Second, Phillips and Sul (2007) show that in a dynamic panel data model, cross-sectional dependence introduces additional small sample bias.

3.2.1 CCEP estimator

Pesaran (2006) proposes to eliminate the error cross-sectional dependence by projecting out the factors f_t using the cross-sectional averages of y_{it}, $y_{i,t-1}$ and x_{it}. For a model with a single factor[9], inserting (30) in (25) and taking cross-sectional averages yields

$$\overline{y}_t = \overline{\alpha} + \rho \overline{y}_{t-1} + \beta' \overline{x}_t + \overline{\gamma} f_t + \phi(L) \overline{\varepsilon}_t, \tag{31}$$

where $\overline{y}_t = N^{-1} \sum_{i=1}^{N} y_{it}$ and similarly for the other variables. Solving (31) for f_t

$$f_t = \frac{1}{\overline{\gamma}} \left(\overline{y}_t - \overline{\alpha} - \rho \overline{y}_{t-1} - \beta' \overline{x}_t - \phi(L) \overline{\varepsilon}_t \right), \tag{32}$$

and inserting (32) in (25) with error structure (30) yields

$$\begin{aligned} y_{it} &= \alpha_i + \rho y_{i,t-1} + \beta' x_{it} + \frac{\gamma_i}{\overline{\gamma}} \left(\overline{y}_t - \overline{\alpha} - \rho \overline{y}_{t-1} - \beta' \overline{x}_t - \phi(L) \overline{\varepsilon}_t \right) + \phi(L) \varepsilon_{it}, \\ &= \alpha_i^+ + \rho y_{i,t-1} + \beta' x_{it} + c_{1i} \overline{y}_t + c_{2i} \overline{y}_{t-1} + c_{3i}' \overline{x}_t + \phi(L) \varepsilon_{it}^+, \end{aligned} \tag{33}$$

with $\alpha_i^+ = \alpha_i - \frac{\gamma_i}{\overline{\gamma}} \overline{\alpha}$, $c_{1i} = \frac{\gamma_i}{\overline{\gamma}}$, $c_{2i} = -\rho \frac{\gamma_i}{\overline{\gamma}}$, $c_{3i} = -\beta' \frac{\gamma_i}{\overline{\gamma}}$, $\varepsilon_{it}^+ = \varepsilon_{it} - \frac{\gamma_i}{\overline{\gamma}} \overline{\varepsilon}_t$.

The CCEP estimator proposed by Pesaran (2006) is the WG estimator applied to the augmented regression in (33). As **A1** implies that $\operatorname{plim}_{N \to \infty} \overline{\varepsilon}_t = 0$, the error made when approximating f_t by $\overline{y}_t$, $\overline{y}_{t-1}$ and $\overline{x}_t$ in (32) becomes negligible for $N \to \infty$ such that $\varepsilon_{it}^+ \xrightarrow{p} \varepsilon_{it}$ in (33). This is the basic result in Pesaran (2006) that the inclusion of cross-sectional averages asymptotically eliminates the error cross-sectional dependence induced by the unobserved common factors. As such, for $N \to \infty$ (33) is a standard dynamic panel data model with individual effects and cross-sectional independent error terms. Similarly to the WG estimator outlined in section 3.1.1, the CCEP estimator therefore exhibits a dynamic panel data bias for fixed T, which disappears for $T \to \infty$. Conditional on x_{it} being predetermined or exogenous and $\phi(L) = 1$, this implies that consistency of the CCEP estimator in a dynamic panel data model requires both N and $T \to \infty$. However, endogeneity of x_{2it} and $MA(q)$ errors μ_{it} imply that the CCEP estimator is inconsistent even for both N and $T \to \infty$. Therefore, we estimate (33) using IV methods in the next section.

3.2.2 CCEP-GMM estimator

For notational convenience, note that the CCEP estimator can also be obtained as the pooled OLS estimator after projecting out the individual effects and the cross-sectional means from the model in

[9] Multiple factors can be treated in the same way (see Phillips and Sul, 2007), and yield the same (unrestricted) model as the one presented in (33) below, but are not presented here for notational convenience.

(25), i.e.

$$\breve{y}_{it} = \rho \breve{y}_{i,t-1} + \beta' \breve{x}_{it} + \breve{\mu}_{it}, \tag{34}$$

where $\breve{y}_{it}$ is the residual from a country-by-country regression of y_{it} on a constant, $\overline{y}_t$, $\overline{y}_{t-1}$ and $\overline{x}_t$ and similarly for the other variables. Letting both N and $T \to \infty$ we have that $\breve{\mu}_{it} \xrightarrow{p} \phi(L)\varepsilon_{it}$ where $N \to \infty$ is required for the elimination of the country-specific effects of the unobserved common factors using the cross-sectional averages and $T \to \infty$ is required to avoid correlation between within transformed variables and errors[10].

Provided both N and T are sufficiently large, such that $\breve{\mu}_{it} \xrightarrow{p} \phi(L)\varepsilon_{it}$, valid moment conditions are

$$E(\breve{y}_{i,t-s}\breve{\mu}_{it}) = 0 \quad \text{for each} \quad q+1 \leq s \leq L+q, \tag{35a}$$

$$E(\breve{x}_{1i,t-s}\breve{\mu}_{it}) = 0 \quad \text{for each} \quad q \leq s \leq L+q-1, \tag{35b}$$

$$E(\breve{x}_{2i,t-s}\breve{\mu}_{it}) = 0 \quad \text{for each} \quad q+1 \leq s \leq L+q. \tag{35c}$$

where in line with the discussion in section 3.1.2 we only use a reduced set of instruments as we have truncated the set of available instruments at the first L available lags and combine all moment conditions for a given instrument lag distance by summing them over T .

3.3 Monte Carlo simulation

This section provides Monte Carlo evidence on the small sample properties of the WG, WG-GMM, CCEP and CCEP-GMM estimators under both cross-sectional dependence and endogeneity. Although we are mainly interested in the setting $T = 35$ and $N = 15$, we also present results for a range of alternative sample sizes to illustrate the more general properties of the estimators.

3.3.1 Experimental design

The data generating process is given by

$$y_{it} = \alpha_i + \rho y_{i,t-1} + \beta x_{it} + \mu_{it}, \qquad \mu_{it} = \gamma_i f_t + \varepsilon_{it}, \tag{36}$$

$$x_{it} = \theta x_{i,t-1} + \lambda_i f_t + \varphi \varepsilon_{it} + \nu_{it}, \tag{37}$$

$$f_t = \tau f_{t-1} + \eta_t, \tag{38}$$

[10]Note that in (34) the individual effects have been removed using a within transformation as the inclusion of a country specific constant in the construction of the de-factored variables implies that, as in (27), these variables are all transformed into deviations from individual means.

where in each replication α_i, ε_{it}, ν_{it} and η_t are all drawn from $i.i.d.N(0,1)$.[11] The model for the individual specific regressors in equation (37) is fairly general as they are allowed to be correlated with the unobserved factors (i.e. $\lambda_i \neq 0$) and to be endogenous (i.e. $\varphi \neq 0$).

We fix $\rho = 0.25$, $\beta = 1$, $\theta = 0.5$ and $\tau = 0.5$[12] and conduct three different experiments:

- Experiment 1: no cross-sectional dependence ($\gamma_i = \lambda_i = 0$) and no endogeneity ($\varphi = 0$)
- Experiment 2: cross-sectional dependence ($\gamma_i \sim i.i.d.U(1,4)$ and $\lambda_i \sim i.i.d.U(1,4)$) and no endogeneity ($\varphi = 0$)
- Experiment 3: cross-sectional dependence ($\gamma_i \sim i.i.d.U(1,4)$ and $\lambda_i \sim i.i.d.U(1,4)$) and endogeneity ($\varphi = 0.5$)

The initial value of y_{it} is set equal to zero and the first 50 observations are discarded before choosing our sample. Each experiment was replicated 2000 times for the (T, N) pairs with $T = 20, 35, 50$ and $N = 15, 50$. The GMM estimators are one-step estimators using $y_{i,t-1}$ and $x_{i,t-1}$ as instruments.

3.3.2 Results

Table B-1 in Appendix B reports mean bias (bias), estimated standard deviation (stde), standard deviation (stdv) and root mean squared error (rmse) in estimating ρ and β. The first experiment is a standard dynamic panel data model with no cross-sectional dependence and no endogeneity. In line with the results in Judson and Owen (1999), the bias of the WG estimator is negligibly small for $T = 35$. The CCEP only has a slightly higher bias and dispersion. As can be expected, the GMM estimators are relatively more biased and have a larger dispersion but are consistent for $T \to \infty$. The second experiment adds cross-sectional dependence, with the unobserved factor being correlated with x_{it}. Both the WG and the WG-GMM estimator now exhibit a considerable omitted variable bias which does not disappear as $T \to \infty$, $N \to \infty$ or both. Moreover the stde considerably underestimates the true stdv. The CCEP estimator is now preferred with the bias being negligibly small for $T = 35$. Again, the CCEP-GMM estimator is relatively more biased and has a larger dispersion for small values of T but is consistent for $T \to \infty$. Estimated standard deviations are fairly accurate for both estimators. The third experiment adds endogeneity. This implies that also the CCEP is inconsistent. The CCEP-GMM estimator is biased for small values of T but this bias disappears as $T \to \infty$. Estimated standard deviations are fairly accurate for moderate values of T. Important to note is that, over the 3 experiments, the bias of the

[11] Note that it is not necessary to control the relative impact of the two error components α_i and μ_{it}, as in Sarafidis and Robertson (2009), since the considered estimators are all of the WG type such that they are invariant to the ratio $\sigma_\alpha^2 / \sigma_\mu^2$.

[12] Results for alternative parameter values are available from the authors on request.

CCEP-type estimators is highly similar for $N = 15$ and $N = 50$. This suggests that a relatively low cross-sectional dimension ($N = 15$) is not really a source of concern.

To summarize, in a dynamic panel data model with both endogeneity and error cross-sectional dependence the CCEP-GMM is the preferred estimator, both in terms of bias of the coefficients and of the estimated standard deviations. Especially when compared to the alternative estimators, it performs reasonably well for the modest sample size $T = 35, N = 15$ that is available for the empirical analysis presented in section 4.

4 Empirical results

Table 2 reports the results from estimating eq.(25) using aggregate yearly data from 15 OECD countries over the period 1972-2007 with the WG, WG-GMM, CCEP, and CCEP-GMM estimators discussed in section 3. The data are described in Appendix C. The CCEP-GMM estimator is our preferred estimator since it corrects for both endogeneity and cross-sectional dependence among the countries in the panel. The latter correction is necessary because (i) table 1 shows that all variables exhibit moderate to strong correlation over countries and (ii) the CD tests for cross-sectional independence reported in table 2 point toward cross-sectional dependence in the errors of the WG-type estimators. Further, as discussed in our Monte Carlo simulations in section 3.3, CCEP-GMM performs reasonably well in samples of modest size.

Table 1: Diagnostic tests for cross-sectional independence

Sample period: 1972-2007, 15 countries

	$\Delta \ln C_{it}$	$\Delta \ln H_{it}$	$\Delta \ln G_{it}$	R_{it}	$(Y_{i,t-1}/C_{i,t-1})$	$(Y_{i,t-1}/C_{i,t-1})^2$	$\Delta \ln Y_{it}$
$\overline{\widehat{\rho}}$	0.274	0.251	0.278	0.678	0.620	0.618	0.288

Note: The average cross correlation coefficient $\overline{\widehat{\rho}} = (2/N(N-1)) \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \widehat{\rho}_{ij}$ is the simple average of the pair-wise cross correlation coefficients $\widehat{\rho}_{ij}$.

The instrument sets used for the GMM estimators are determined by setting $q = 1$ and $L = 3$. From the Sargan/Hansen overidentifying restrictions test S_1 reported in table 2 we note that the implied moment conditions are not rejected by the data. This contrasts with the case $q = 0$ and $L = 4$ for which the moment conditions are rejected by the Sargan/Hansen test S_0. These results suggest that there might be serial correlation of order 1 in the error term. This is confirmed by the Difference-in-Sargan test ΔS_1 which rejects the null of no serial correlation of order 1 for all estimators. When we look at the Difference-in-Sargan test ΔS_2 which tests the null of no serial correlation of order 2 we find that the null is still rejected at conventional levels of significance for the WG-GMM estimators. It is not rejected for the CCEP-GMM estimator however. This further contributes to its status as our preferred estimator.

From table 2 we first note that the coefficient on lagged aggregate consumption growth is positive

Table 2: Alternative panel data estimation results using labour income

Dependent variable: $\Delta \ln C_{it}$ Sample period: 1972-2007, 15 countries

	WG	WG-GMM		CCEP	CCEP-GMM	
		one-step	two-step		one-step	two-step
Coefficient estimates ($q = 1, L = 3$)						
$\Delta \ln C_{i,t-1}$	0.00	0.14	0.10	0.05	-0.17	-0.21^{*}
	(0.04)	(0.17)	(0.15)	(0.05)	(0.13)	(0.12)
$\Delta \ln H_{it}$	0.04	0.59^{***}	0.61^{***}	0.04	0.52^{***}	0.49^{***}
	(0.04)	(0.15)	(0.14)	(0.04)	(0.14)	(0.13)
$\Delta \ln G_{it}$	0.33^{***}	0.28	0.29^{*}	0.30^{***}	0.17	0.21
	(0.03)	(0.17)	(0.16)	(0.04)	(0.16)	(0.15)
R_{it}	0.01	0.05	0.08^{*}	-0.01	-0.15	-0.22
	(0.03)	(0.04)	(0.04)	(0.03)	(0.21)	(0.19)
$(Y_{i,t-1}/C_{i,t-1})$	0.18	0.08	0.10	-0.10	-2.09^{*}	-2.66^{**}
	(0.16)	(0.19)	(0.19)	(0.46)	(1.08)	(1.05)
$(Y_{i,t-1}/C_{i,t-1})^2$	-0.05	-0.02	-0.02	0.22	1.39^{**}	1.74^{***}
	(0.08)	(0.10)	(0.09)	(0.24)	(0.61)	(0.59)
$\Delta \ln Y_{it}$	0.03	0.17	0.16	0.05	0.52^{***}	0.72^{***}
	(0.02)	(0.27)	(0.23)	(0.04)	(0.18)	(0.16)
$\Delta \ln Y_{i,t-1}$	0.50^{***}	-0.31^{**}	-0.28^{**}	0.47^{***}	0.13	0.16
	(0.04)	(0.15)	(0.14)	(0.04)	(0.12)	(0.11)
Residual cross-sectional independence tests						
$\overline{\widehat{\rho}}$	0.18	0.10	0.10	-0.06	-0.05	-0.03
CD	10.93	6.16	5.98	-3.38	-2.87	-1.52
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.13]
Overidentifying restrictions tests						
Sargan/Hansen						
S_0 ($q = 0, L = 4, df = 21$)		45.54	36.73		35.09	32.75
		[0.00]	[0.02]		[0.03]	[0.05]
S_1 ($q = 1, L = 3, df = 14$)		20.18	17.59		18.68	19.42
		[0.12]	[0.23]		[0.18]	[0.15]
S_2 ($q = 2, L = 2, df = 7$)		2.22	1.89		12.09	10.49
		[0.95]	[0.97]		[0.10]	[0.16]
Difference-in-Sargan						
$\Delta S_1 = S_0 - S_1$ ($df = 7$)		25.36	19.14		16.41	13.33
		[0.00]	[0.01]		[0.02]	[0.06]
$\Delta S_2 = S_1 - S_2$ ($df = 7$)		17.96	15.7		6.59	8.93
		[0.01]	[0.03]		[0.47]	[0.26]

Notes: Standard errors are in parentheses, p-values are in square brackets. One-step GMM uses the 'two stage least squares' suboptimal choice of weighting matrix while two-step GMM uses a consistent estimate for the optimal weighting matrix constructed from a Newey-West type of estimator allowing for general cross-section and time series heteroscedasticity and $MA(1)$ errors. Both one-step and two-step GMM standard errors are robust to general cross-section and time series heteroscedasticity and to $MA(1)$ errors. *, ** and *** indicate significance at the 10%, 5% and 1% level respectively.

The average cross correlation coefficient $\overline{\widehat{\rho}}$ is defined in Table 1. CD is the Pesaran (2004) test defined as $\sqrt{2T/N(N-1)}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\widehat{\rho}_{ij}$, which is asymptotically normal under the null of cross-sectional independence. The Sargan/Hansen test of overidentifying restrictions is χ^2 distributed under to null of joint validity of all moment conditions defined in equations (29) or (35) for given choice of q and L. The difference-in-Sargan test is χ^2 distributed under to null of joint validity of a subset of moment conditions, i.e. ΔS_1 under the null of no first order serial correlation in the error terms and ΔS_2 under the null of no second order serial correlation in the error terms.

only if we do not model cross-sectional dependence (WG-type estimators) or do not use instruments (CCEP estimator). When we model cross-sectional dependence and correct for endogeneity we find that the coefficient on lagged aggregate consumption growth turns negative, albeit it is either insignificant or its significance is quite low. This result is in accordance with Kiley (2007) who finds generally negative and insignificant estimates for the habit formation parameter. This result differs from the conclusions reported by Carroll et al. (2008) who generally find positive and significant values for this parameter.

We further find that the impact of aggregate disposable income growth on aggregate consumption growth is positive and strongly significant for the WG estimator and for the CCEP and CCEP-GMM estimators. Estimated with our preferred CCEP-GMM estimator, the magnitude of the estimated coefficient on disposable income growth is 0.52 for the one-step CCEP-GMM estimator and 0.72 for the two-step CCEP-GMM estimator. The former estimate is in accordance with the magnitude of the estimates found for this coefficient in other studies (see e.g. Campbell and Mankiw, 1989, 1990, 1991).

Contrary to Carroll et al. (2008) our results thus favour the Campbell and Mankiw model over the model with habit formation. One reason may be that our specification is more elaborate since we include other variables like the growth rate in hours worked which is not included in Carroll et al.[13]. In particular, we find that this variable is positive and highly significant across all GMM estimators. In fact, the growth rate in hours worked turns out to be the most robust explanatory variable for aggregate consumption growth. This provides support for the analysis of Basu and Kimball (2002) who argue in favour of complementarity between consumption and labour in the US. Kiley (2007) finds, also for the US, that hours worked explain aggregate consumption growth better than either lagged consumption growth or disposable income growth. Hence our analysis seems to confirm this complementarity hypothesis between consumption and labour in a panel of OECD countries.

We then look at potential intertemporal substitution effects, i.e. the effects of the real interest rate on aggregate consumption growth. Our results are fully in line with the literature in the sense that the real interest rate has an insignificant impact on aggregate consumption growth in almost all cases (see e.g. Campbell and Mankiw, 1990). The exception is the result obtained with the two-step WG-GMM estimator where the impact of the real interest rate is positive and significant. While this seems to suggest that there are intertemporal substitution effects of modest magnitude in our sample, we do not favour these results. The reason is that the WG-GMM estimator does not take into account cross-sectional dependence in the panel. When estimating the effect of the real interest rate on aggregate consumption growth with the CCEP-GMM estimator instead we find no evidence of intertemporal substitution effects.

The impact of government consumption growth on private consumption growth is less clear-cut. We

[13] Another reason may be that we use aggregate total consumption instead of consumption of non-durables and services (which Carroll et al. use for about half their countries). Durable components in our consumption measure may bias our habit parameter estimate downward since durable consumption growth is somewhat negatively autocorrelated (see Mankiw, 1982). Kiley (2007) reports negative and insignificant estimates for the habit parameter both when total consumption and non-durable consumption and services are used to measure consumption. In the latter case the estimates are less negative however.

find that the estimates of this impact are positive and significant when estimated using WG or CCEP but insignificant when estimated with the GMM estimators. The only exception is again the result obtained with the two-step WG-GMM estimator. Hence, we don't find strong support for the complementarity between private and public consumption as reported for instance by Evans and Karras (1998).

Evidence of precautionary savings effects on aggregate consumption growth can be obtained by looking at the estimate for the coefficient on the squared income to consumption ratio. We find that this coefficient is insignificant for all estimators but the CCEP-GMM estimator. When using our preferred estimator this coefficient has the expected positive sign and is highly significant. The CCEP-GMM estimator thus facilitates the detection of significant precautionary effects at the aggregate level in a panel of OECD countries. This result adds to the literature that finds evidence that precautionary savings matter for aggregate consumption growth (see e.g. Parker and Preston, 2005).

To summarize, our newly introduced CCEP-GMM estimator - which corrects for endogeneity and error cross-sectional dependence and is robust to heteroscedasticity and $MA(1)$ serial correlation - indicates that aggregate consumption growth depends positively on aggregate hours worked (hence supporting the labour-consumption complementarity hypothesis), depends positively on aggregate disposable income growth (hence supporting the rule-of-thumb consumption model of Campbell and Mankiw), and depends positively on the squared income to consumption ratio (hence supporting the notion that precautionary savings matter for the aggregate economy).

5 Conclusions

This paper examines the sources of stickiness in aggregate private consumption growth. We first derive a dynamic consumption equation which nests recent developments in consumption theory: rule-of-thumb consumption, habit formation, non-separabilities between both private consumption and hours worked and private consumption and government consumption, intertemporal substitution effects and precautionary savings. Next, we estimate this dynamic consumption equation for a panel of 15 OECD countries over the period 1972-2007. We follow recent developments in panel data econometrics by allowing for unobserved common factors which have heterogeneous impacts on the countries in the panel. We develop a CCEP-GMM estimator by combining the CCEP estimator advanced by Pesaran (2006) to account for error cross-sectional dependence and the GMM estimator to account for endogeneity of the regressors. We show that the moment conditions imposed by this CCEP-GMM estimator are valid as $N, T \rightarrow \infty$ jointly. A Monte Carlo experiment shows that the CCEP-GMM estimator performs reasonably well even for the modest sample size $T = 35, N = 15$ that is available for our empirical analysis. In our dynamic panel data setting with both endogeneity and error cross-sectional dependence, it is preferred over standard WG, WG-GMM and CCEP estimators both in terms of bias of the coefficients and in terms of the estimated standard deviations.

Taking into account endogeneity and cross-sectional dependence proves to be important as it has a marked effect on our estimation results. These suggest that the growth rate in aggregate private consumption depends positively on aggregate hours worked. This supports the labour-consumption complementarity hypothesis of Basu and Kimball (2002). It also depends positively on aggregate disposable income growth. This supports the rule-of-thumb consumption model of Campbell and Mankiw (1989, 1990, 1991). Finally it depends positively on the squared income to consumption ratio. This supports the notion that precautionary savings matter for aggregate consumption growth (see Parker and Preston, 2005). We find little or no support for habit formation, non-separabilities between private consumption and government consumption and intertemporal substitution effects as lagged consumption growth, government consumption and the real interest rate are generally found to be insignificant.

References

Arellano, M., Bond, S., 1991. Some tests of specification for panel data: Monte carlo evidence and anapplication to employment equations. Review of Economic Studies 58, 277–297.

Banks, J., Blundell, R., Brugiavini, A., 2001. Risk pooling, precautionary saving and consumption growth. Review of Economic Studies 68, 757–779.

Basu, S., Kimball, M., 2002. Long-run labor supply and the elasticity of intertemporal substitution for consumption, mimeo.

Blundell, R., Bond, S., 1998. Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics 87, 115–143.

Bun, M., Windmeijer, F., 2007. The weak instrument problem of the system gmm estimator in dynamic panel data models. UvA-Econometrics Discussion Paper 2007/01, Amsterdam School of Economics.

Campbell, J., 1998. Asset prices, consumption, and the business cycle. NBER Working Paper (6485).

Campbell, J., Mankiw, N., 1989. Consumption, income, and interest rates: reinterpreting the time seriesevidence. NBER Macroeconomics Annual.

Campbell, J., Mankiw, N., 1990. Permanent income, current income and consumption. Journal of Business and Economic Statistics 8, 265–79.

Campbell, J., Mankiw, N., 1991. The response of consumption to income. a cross country investigation. European Economic Review 35, 723–67.

Carey, D., Rabesona, J., 2004. Tax ratios on labour and capital income and on consumption. In: Sorensen, P. (Ed.), Measuring the tax burden on labour and capital. MIT Press.

Carroll, C., Slacalek, J., Sommer, M., 2008. International evidence on sticky consumption growth. ECB Working Paper Series (886).

Coakley, J., Fuertes, A., Smith, R., 2002. A principal components approach to cross-section dependence in panels, birkbeck College, University of London.

Coakley, J., Fuertes, A., Smith, R., 2006. Unobserved heterogeneity in panel time series models. Computational Statistics & Data Analysis 50 (9), 2361–2380.

Deaton, A., 1991. Savings and liquidity constraints. Econometrica 59, 1,221–48.

Evans, P., Karras, G., 1998. Liquidity constraints and the substitutability between private and governmentconsumption. Economic Inquiry 36, 203–14.

Fiorito, R., Padrini, F., 2001. Distortionary taxation and labour market performance. Oxford Bulletin of Economics and Statistics 63 (2), 173–96.

Flavin, M., 1985. Excess sensitivity of consumption to current income: liquidity constraints or myopia? Canadian Journal of Economics 38, 117–36.

Goodfriend, M., 1992. Information-aggregation bias. American Economic Review 82, 508–519.

Gourinchas, P., Parker, J., 2001. The empirical importance of precautionary saving. American Economic Review 91 (2), 406–12.

Hall, R., 1978. Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence. Journal of Political Economy 86 (6), 971–87.

Holtz-Eakin, D., Newey, W., Rosen, H., 1988. Estimating vector autoregressions with panel data. Econometrica 56 (6), 1371–1395.

Jappelli, T., Pagano, M., 1989. Consumption and capital market imperfections: an international comparison. American Economic Review 79 (5), 1,088–105.

Judson, A., Owen, A., 1999. Estimating dynamic panel data models: A guide for macroeconomists. Economics Letters 65, 9–15.

Kiley, M., 2007. Habit persistence, non-separability between consumption and leisure, or rule-of-thumb consumers: which accounts for the predictability of consumptiongrowth?, mimeo.

Kiviet, J., 1995. On bias, inconsistency, and efficiency of various estimators in dynamicpanel data models. Journal of Econometrics 68, 53–78.

Lopez, H., Schmidt-Hebbel, K., Servén, L., 2000. How effective is fiscal policy in raising national saving? Review of Economics and Statistics 82 (2), 226–238.

Mankiw, N., 1982. Hall's consumption hypothesis and durable goods. Journal of Monetary Economics 10, 417–26.

Meghir, C., Pistaferri, L., 2004. Income variance dynamics and heterogeneity. Econometrica 72 (1), 1–32.

Nickell, S., 1981. Biases in dynamic models with fixed effects. Econometrica 49 (6), 1417–1426.

Parker, J., Preston, B., 2005. Precautionary savings and consumption fluctuations. American Economic Review 95 (4), 1119–1143.

Pesaran, M., 2004. General diagnostic tests for cross section dependence in panels. Cambridge Working Papers in Economics 0435, Faculty of Economics, University of Cambridge.

Pesaran, M., 2006. Estimation and inference in large heterogeneous panels with a multifactor error structure. Econometrica 74 (4), 967–1012.

Phillips, P. C. B., Sul, D., 2003. Dynamic panel estimation and homogeneity testing under cross section dependence. Econometrics Journal 6 (1), 217–259.

Phillips, P. C. B., Sul, D., 2007. Bias in dynamic panel estimation with fixed effects, incidental trends and cross section dependence. Journal of Econometrics 137 (1), 162–188.

Pischke, J., 1995. Individual income, incomplete information and aggregate consumption. Econometrica 63 (4), 805–40.

Pozzi, L., Malengier, G., 2007. Certainty equivalence and the excess sensitivity of private consumption. Journal of Money Credit and Banking 39 (7), 1839–1848.

Roodman, D., 2009. A note on the theme of too many instruments. Oxford Bulletin of Economics and Statistics 71 (1), 135–158.

Sarafidis, V., Robertson, D., 2009. On the impact of error cross-sectional dependence in short dynamic panel estimation. Econometrics Journal 12 (1), 62–81.

Tauchen, G., 1986. Statistical properties of gmm estimators of structural parameters obtained from financial market data. Journal of Business and Economic Statistics 4, 397–416.

Westerlund, J., 2008. Panel cointegration tests of the fisher effect. Journal of Applied Econometrics 23 (2), 193–233.

Working, H., 1960. Note on the correlation of first differences of averages in a random chain. Econometrica 28, 916–18.

Ziliak, J., 1997. Efficient estimation with panel data when instruments are predetermined: an empirical comparison of moment-condition estimators. Journal of Business and Economic Statistics 14 (4), 419–431.

Appendix A Derivations

A.1 Approximating $E_{t-1}\varepsilon_t$ and $V_{t-1}\varepsilon_t$

First, a second order Taylor expansion of a function $f(x)$ around $x = x_0$ is given by $f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2$. Now for $f(x) = \varepsilon_t = \exp(\ln \varepsilon_t)$, $x = \ln \varepsilon_t$, and $x_0 = E_{t-1} \ln \varepsilon_t = 0$ we find that $\varepsilon_t = 1 + \ln \varepsilon_t + \frac{1}{2}(\ln \varepsilon_t)^2$ so that, with the use of eqs. (3) and (4), we can write $E_{t-1}\varepsilon_t = 1 + \frac{1}{2}\sigma^2_{\ln \varepsilon} = z$ (with $z > 1$).

Second, note that $V_{t-1}\varepsilon_t = E_{t-1}\varepsilon_t^2 - z^2$. Now for $f(x) = \varepsilon_t^2 = \exp(2 \ln \varepsilon_t)$, $x = \ln \varepsilon_t$, and $x_0 = E_{t-1} \ln \varepsilon_t = 0$ we can use a second order Taylor expansion of $f(x)$ around $x = x_0$ to calculate $\varepsilon_t^2 = 1 + 2\ln(\varepsilon_t) + 2(\ln \varepsilon_t)^2$ so that $E_{t-1}\varepsilon_t^2 = 1 + 2\sigma^2_{\ln \varepsilon}$. Now we have $V_{t-1}\varepsilon_t = E_{t-1}\varepsilon_t^2 - z^2 = (1 + 2\sigma^2_{\ln \varepsilon}) - (1 + \sigma^2_{\ln \varepsilon} + \frac{1}{4}(\sigma^2_{\ln \varepsilon})^2) = \sigma^2_{\ln \varepsilon} - \frac{1}{4}(\sigma^2_{\ln \varepsilon})^2 \approx \sigma^2_{\ln \varepsilon}$.

A.2 Approximating the function $f(w_t, \ln h_t, \ln G_t)$ around certainty

A second order Taylor expansion of a function $f(x, y, z)$ around $x = x_0$, $y = y_0$, and $z = z_0$ is given by $f(x, y, z) = f(x_0, y_0, z_0) + \sum_i \frac{\partial f}{\partial i}(x_0, y_0, z_0)(i - i_0) + \frac{1}{2}\sum_i \sum_j \frac{\partial^2 f}{\partial i \partial j}(x_0, y_0, z_0)(i - i_0)(j - j_0)$ for $i = x, y, z$ and $j = x, y, z$. With this formula we approximate $f(w_t, \ln h_t, \ln G_t) = (\alpha w_t)^{-\theta} e^{\gamma \ln h_t} e^{\pi \ln G_t}$around $w_t = \overline{w}_t$, around $\ln h_t = \overline{\ln h}_t$ (where $\overline{\ln h}_t = E_{t-1} \ln h_t$), and around $\ln G_t = \overline{\ln G}_t$ (where $\overline{\ln G}_t = E_{t-1} \ln G_t$) by calculating the necessary derivatives.

A.3 Calculating aggregate consumption growth $\Delta \ln(C_t)$

Define total consumption in the economy C_t as

$$C_t = C_{1,t} + C_{2,t}. \tag{A-1}$$

Following Campbell and Mankiw (1990) we assume that the labour income of type 2 consumers $Y_{2,t}$ is a fixed fraction $(1 - \lambda)$ of total income Y_t where $0 < \lambda \leq 1$. Hence $Y_{2,t} = (1 - \lambda)Y_t$. Since $Y_t = Y_{1,t} + Y_{2,t}$ this implies $Y_{1,t} = \lambda Y_t$. Since $C_{2,t} = Y_{2,t}$ this gives,

$$C_{2,t} = (1 - \lambda)Y_t.$$

Eq.(A-1) is now given by

$$C_t = C_{1,t} + (1 - \lambda)Y_t. \tag{A-2}$$

We now take the log of C_t. Following Campbell and Mankiw (1990, p.268) we approximate the log of an average by an average of logs $\ln(ax+bz) \simeq a\ln(x) + b\ln(z)$ so that from eq.(A-2) we have

$$\ln(C_t) \simeq \ln(C_{1,t}) + (1-\lambda)\ln(Y_t). \tag{A-3}$$

In first differences this gives eq.(18) in the main text.

A.4 Replacing the ratios $Y_{1,t}/C_t^1$ and $(Y_{1,t}/C_{1,t})^2$

We replace $Y_{1,t}/C_{1,t}$ and $(Y_{1,t}/C_{1,t})^2$ in eq.(16) with their observable counterparts Y_t/C_t and $(Y_t/C_t)^2$.

First, note that $C_{1,t} = C_t - (1-\lambda)Y_t$ and $Y_{1,t} = \lambda Y_t$. Hence, $C_{1,t}/Y_{1,t} = (1/\lambda)(C_t/Y_t) - (1-\lambda)/\lambda = a + b(C_t/Y_t)$ where $b = 1/\lambda > 0$ and $a = -(1-\lambda)/\lambda < 0$. We have $a+b=1$. Now from this we want an expression for $Y_{1,t}/C_{1,t}$. Note that we can write $Y_{1,t}/C_{1,t} = (a + b(C_t/Y_t))^{-1}$. We write this expression as $f(z) = (a+bz^{-1})^{-1}$ where $f(z) = \frac{Y_{1,t}}{C_{1,t}}$, and where $z = \frac{Y_t}{C_t}$. We then linearize this expression with a first-order Taylor expansion of $f(z)$ around $z=1$. This gives $f(z) \simeq (a+b)^{-1} - b(a+b)^{-2} + b(a+b)^{-2}z$ or $f(z) \simeq (1-b) + bz$ since $a+b=1$. Hence we find $(Y_{1,t}/C_{1,t}) = (1-b) + b(Y_t/C_t) = 1 - \lambda^{-1} + \lambda^{-1}(Y_t/C_t)$. Substituting $Y_{1,t}/C_{1,t}$ (lagged once) into eq.(16) gives an expression for aggregate consumption growth in terms of the observable Y_t/C_t.

Second, note that $(C_{1,t})^2 = (C_t - (1-\lambda)Y)^2 = C_t^2 + (1-\lambda)^2 Y_t^2 - 2(1-\lambda)C_tY_t$ and that $(Y_{1,t})^2 = \lambda^2 Y_t^2$. Hence, $(C_{1,t}/Y_{1,t})^2 = (1/\lambda^2)(C_t/Y_t)^2 - (2(1-\lambda)/\lambda^2)(C_t/Y_t) + (1-\lambda)^2/\lambda^2 = a + b(C_t/Y_t) + c(C_t/Y_t)^2$ where $a = (1-\lambda)^2/\lambda^2 > 0$, $b = -(2(1-\lambda)/\lambda^2) < 0$, $c = 1/\lambda^2 > 0$. We have $a+b+c=1$. Now from this we want an expression for $(Y_{1,t}/C_{1,t})^2$. Note that we can write $(Y_{1,t}/C_{1,t})^2 = (a + b(C_t/Y_t) + c(C_t/Y_t)^2)^{-1}$. We write this expression as $f(z) = (a + bz^{-0.5} + cz^{-1})^{-1}$ where $f(z) = (Y_{1,t}/C_{1,t})^2$, and where $z = (Y_t/C_t)^2$. We then take a first-order Taylor expansion of $f(z)$ around $z=1$. This gives $f(z) \simeq (a+b+c)^{-1} - (b/2+c)(a+b+c)^{-2} + (b/2+c)(a+b+c)^{-2}z$ or $f(z) \simeq 1 - (b/2+c) + (b/2+c)z$ since $a+b+c=1$. Hence we find $(Y_{1,t}/C_{1,t})^2 = 1 - (b/2+c) + (b/2+c)\left(Y_t/C_t\right)^2 = 1 - \lambda^{-1} + \lambda^{-1}\left(Y_t/C_t\right)^2$ where the latter equality follows from $b/2 + c = \lambda^{-1}$. Substituting $(Y_{1,t}/C_{1,t})^2$ (lagged once) into eq.(16) gives an expression for aggregate consumption growth in terms of the observable $(Y_t/C_t)^2$.

A.5 Proof of $\lim_{\lambda\to 0}\Delta\ln C_t = \Delta\ln Y_t$ in eq.(24)

We first take the limits of eqs.(21) and (22),

$$\lim_{\lambda\to 0}(Y_{1,t-1}/C_{1,t-1}) = \lim_{\lambda\to 0}\frac{\lambda - 1 + (Y_{t-1}/C_{t-1})}{\lambda} = \frac{0}{0}, \tag{A-4}$$

$$\lim_{\lambda\to 0}(Y_{1,t-1}/C_{1,t-1})^2 = \lim_{\lambda\to 0}\frac{\lambda - 1 + (Y_{t-1}/C_{t-1})^2}{\lambda} = \frac{0}{0}, \tag{A-5}$$

where in the last step we use $\lim_{\lambda\to 0}(Y_{t-1}/C_{t-1}) = 1$ and $\lim_{\lambda\to 0}(Y_{t-1}/C_{t-1})^2 = 1$ as when all income is earned by consumers of type 2 then total consumption in the economy equals total income. With the use of l'Hopital rule we obtain

$$\lim_{\lambda\to 0}(Y_{1,t-1}/C_{1,t-1}) = \lim_{\lambda\to 0}\frac{\frac{d(\lambda-1+(Y_{t-1}/C_{t-1}))}{d\lambda}}{\frac{d\lambda}{d\lambda}} = \lim_{\lambda\to 0}\frac{1+\frac{d(Y_{t-1}/C_{t-1})}{d\lambda}}{1} = 1, \tag{A-6}$$

$$\lim_{\lambda\to 0}(Y_{1,t-1}/C_{1,t-1})^2 = \lim_{\lambda\to 0}\frac{\frac{d(\lambda-1+(Y_{t-1}/C_{t-1})^2)}{d\lambda}}{\frac{d\lambda}{d\lambda}} = \lim_{\lambda\to 0}\frac{1+\frac{d((Y_{t-1}/C_{t-1})^2)}{d\lambda}}{1} = 1, \tag{A-7}$$

where in the last step we use $\lim_{\lambda\to 0}\frac{d(Y_{t-1}/C_{t-1})}{d\lambda} = 0$ and $\lim_{\lambda\to 0}\frac{d((Y_{t-1}/C_{t-1})^2)}{d\lambda} = 0$ as the derivatives of the ratios (Y_{t-1}/C_{t-1}) and $(Y_{t-1}/C_{t-1})^2$ approach zero because when $\lambda \to 0$ the ratios (Y_{t-1}/C_{t-1}) and $(Y_{t-1}/C_{t-1})^2$ approach the constant 1.

We now take the limit of eq.(16) while using eqs.(A-6) and (A-7),

$$\lim_{\lambda\to 0}\Delta\ln C_{1,t} = \left[\begin{array}{c} \left(\frac{k_0-\delta}{\theta}\right) + \frac{\beta(\theta-1)}{\theta}\lim_{\lambda\to 0}\Delta\ln C_{1,t-1} + \frac{\gamma}{\theta}\lim_{\lambda\to 0}\Delta\ln H_{1,t} \\ +\frac{\pi}{\theta}\Delta\ln G_t + \frac{1}{\theta}R_t + \frac{k_1}{\theta} + \frac{k_2}{\theta} + \lim_{\lambda\to 0}\Psi_t \end{array}\right]. \tag{A-8}$$

When $\lambda \to 0$, type 1 consumers have no income hence they do not work hours and they do not consume, i.e $\lim_{\lambda\to 0}\Delta\ln C_{1,t} = 0$, $\lim_{\lambda\to 0}\Delta\ln C_{1,t-1} = 0$, and $\lim_{\lambda\to 0}\Delta\ln H_{1,t} = 0$. Substituting these results into eq.(A-8) gives

$$\left(\frac{k_0-\delta}{\theta}\right) + \frac{k_1}{\theta} + \frac{k_2}{\theta} + \frac{\pi}{\theta}\Delta\ln G_t + \frac{1}{\theta}R_t + \lim_{\lambda\to 0}\Psi_t = 0. \tag{A-9}$$

Note that we can write eq.(24) as,

$$\begin{aligned}\Delta\ln C_t \quad = \quad & \left(\frac{k_0-\delta}{\theta}\right) + \frac{k_1}{\theta}\left(\frac{Y_{1,t-1}}{C_{1,t-1}}\right) + \frac{k_2}{\theta}\left(\frac{Y_{1,t-1}}{C_{1,t-1}}\right)^2 + \frac{\pi}{\theta}\Delta\ln G_t + \frac{1}{\theta}R_t + \Psi_t \\ & +\frac{\beta(\theta-1)}{\theta}\Delta\ln C_{t-1} + \frac{\gamma}{\theta}\Delta\ln H_{1,t} + (1-\lambda)\Delta\ln Y_t - (1-\lambda)\left(\frac{\beta(\theta-1)}{\theta}\right)\Delta\ln Y_{t-1}.\end{aligned} \tag{A-10}$$

Taking the limit of eq.(A-10) while using the result $\lim_{\lambda\to 0}\Delta\ln H_{1,t} = 0$ as well as eqs.(A-6), (A-7) and (A-9) gives

$$\lim_{\lambda\to 0}\Delta\ln C_t - \frac{\beta(\theta-1)}{\theta}\lim_{\lambda\to 0}\Delta\ln C_{t-1} = \lim_{\lambda\to 0}\Delta\ln Y_t - \left(\frac{\beta(\theta-1)}{\theta}\right)\lim_{\lambda\to 0}\Delta\ln Y_{t-1}. \tag{A-11}$$

The last equality can only hold if $\Delta\ln C_t = \Delta\ln Y_t$ and if $\Delta\ln C_{t-1} = \Delta\ln Y_{t-1}$.

Appendix B Results Monte-Carlo simulation

Table B-1: Monte Carlo results ($\rho = 0.25$, $\beta = 1$, $\theta = 0.5$, $\tau = 0.5$)

	WG				CCEP				WG-GMM				CCEP-GMM			
(T, N)	bias	stde	stdv	rmse	bias	stde	stdv	rmse	bias	stde	stdv	rmse	bias	stde	stdv	rmse
Experiment 1: $\gamma_i = \lambda_i = 0$ and $\varphi = 0$																
Results for ρ																
(20,15)	−0.030	0.040	0.039	0.049	−0.050	0.044	0.047	0.068	−0.066	0.060	0.062	0.090	−0.112	0.070	0.084	0.141
(35,15)	−0.016	0.029	0.029	0.033	−0.027	0.031	0.032	0.042	−0.036	0.042	0.044	0.057	−0.060	0.044	0.050	0.078
(50,15)	−0.013	0.024	0.025	0.028	−0.020	0.025	0.027	0.033	−0.026	0.035	0.037	0.046	−0.042	0.036	0.040	0.058
(20,50)	−0.031	0.022	0.022	0.038	−0.052	0.024	0.027	0.059	−0.065	0.031	0.031	0.072	−0.110	0.033	0.044	0.118
(35,50)	−0.017	0.016	0.016	0.023	−0.028	0.017	0.017	0.033	−0.035	0.023	0.023	0.042	−0.059	0.023	0.026	0.064
(50,50)	−0.011	0.013	0.013	0.017	−0.019	0.014	0.014	0.023	−0.024	0.019	0.019	0.031	−0.040	0.019	0.021	0.045
Results for β																
(20,15)	0.013	0.059	0.056	0.058	0.015	0.065	0.067	0.068	0.181	0.220	0.229	0.292	0.364	0.306	0.383	0.529
(35,15)	0.008	0.043	0.043	0.043	0.014	0.045	0.047	0.049	0.093	0.138	0.143	0.171	0.162	0.156	0.176	0.240
(50,15)	0.007	0.035	0.037	0.038	0.010	0.037	0.039	0.041	0.064	0.111	0.119	0.135	0.104	0.120	0.133	0.169
(20,50)	0.014	0.032	0.032	0.035	0.020	0.035	0.037	0.042	0.177	0.110	0.111	0.208	0.338	0.135	0.176	0.381
(35,50)	0.007	0.023	0.024	0.025	0.012	0.025	0.025	0.027	0.086	0.075	0.076	0.114	0.152	0.081	0.091	0.177
(50,50)	0.006	0.019	0.019	0.020	0.009	0.020	0.020	0.022	0.060	0.060	0.061	0.086	0.101	0.063	0.068	0.122
Experiment 2: $\gamma_i = i.i.d.U[1,4]$, $\lambda_i = i.i.d.U[1,4]$ and $\varphi = 0$																
Results for ρ																
(20,15)	−0.155	0.021	0.036	0.160	−0.060	0.043	0.045	0.075	0.177	0.191	0.267	0.320	−0.130	0.105	0.130	0.184
(35,15)	−0.148	0.015	0.031	0.151	−0.032	0.030	0.031	0.045	0.230	0.097	0.144	0.271	−0.066	0.042	0.047	0.081

Continued on next page

Continued from previous page

	WG				CCEP				WG-GMM				CCEP-GMM			
(T,N)	bias	stde	stdv	rmse	bias	stde	stdv	rmse	bias	stde	stdv	rmse	bias	stde	stdv	rmse
(50,15)	−0.148	0.013	0.028	0.151	−0.023	0.024	0.027	0.035	0.241	0.080	0.119	0.269	−0.046	0.034	0.039	0.060
(20,50)	−0.153	0.012	0.036	0.157	−0.060	0.024	0.027	0.066	0.170	0.064	0.165	0.237	−0.119	0.032	0.043	0.127
(35,50)	−0.146	0.008	0.027	0.148	−0.033	0.016	0.017	0.037	0.234	0.050	0.122	0.263	−0.065	0.022	0.024	0.069
(50,50)	−0.145	0.007	0.022	0.147	−0.022	0.013	0.014	0.026	0.247	0.042	0.101	0.267	−0.044	0.018	0.020	0.048
Results for β																
(20,15)	0.728	0.040	0.102	0.735	0.019	0.066	0.066	0.068	−0.244	0.598	0.722	0.761	0.485	0.613	0.845	0.974
(35,15)	0.720	0.030	0.094	0.726	0.016	0.045	0.047	0.050	−0.313	0.266	0.285	0.423	0.187	0.159	0.180	0.259
(50,15)	0.723	0.024	0.093	0.729	0.011	0.037	0.040	0.041	−0.325	0.212	0.233	0.400	0.116	0.119	0.131	0.175
(20,50)	0.716	0.022	0.070	0.720	0.022	0.036	0.036	0.043	−0.218	0.191	0.269	0.346	0.412	0.152	0.213	0.464
(35,50)	0.715	0.016	0.061	0.718	0.014	0.025	0.025	0.029	−0.312	0.134	0.188	0.364	0.177	0.086	0.089	0.198
(50,50)	0.716	0.014	0.056	0.718	0.011	0.020	0.020	0.022	−0.331	0.110	0.155	0.365	0.114	0.063	0.067	0.132
Experiment 3: $\gamma_i = i.i.d.U[1,4]$, $\lambda_i = i.i.d.U[1,4]$ and $\varphi = 0.5$																
Results for ρ																
(20,15)	−0.171	0.018	0.032	0.174	−0.131	0.032	0.033	0.135	0.200	0.827	0.720	0.747	−0.126	0.518	0.261	0.290
(35,15)	−0.165	0.014	0.028	0.167	−0.116	0.022	0.024	0.119	0.262	0.116	0.159	0.306	−0.066	0.046	0.054	0.085
(50,15)	−0.165	0.011	0.025	0.166	−0.110	0.018	0.020	0.112	0.273	0.093	0.132	0.304	−0.045	0.037	0.042	0.061
(20,50)	−0.169	0.010	0.031	0.172	−0.132	0.017	0.019	0.133	0.197	0.058	0.181	0.267	−0.124	0.031	0.042	0.131
(35,50)	−0.163	0.008	0.024	0.164	−0.116	0.012	0.012	0.117	0.262	0.057	0.130	0.293	−0.067	0.023	0.025	0.072
(50,50)	−0.162	0.006	0.019	0.163	−0.110	0.010	0.011	0.110	0.276	0.048	0.108	0.296	−0.046	0.019	0.021	0.050
Results for β																
(20,15)	0.771	0.036	0.097	0.777	0.395	0.053	0.057	0.399	−0.326	2.853	2.433	2.453	0.364	3.171	1.568	1.609

Continued on next page

Continued from previous page

	WG				CCEP				WG-GMM				CCEP-GMM			
(T, N)	bias	stde	stdv	rmse	bias	stde	stdv	rmse	bias	stde	stdv	rmse	bias	stde	stdv	rmse
(35,15)	0.763	0.027	0.092	0.769	0.386	0.037	0.042	0.389	−0.413	0.320	0.330	0.529	0.148	0.191	0.226	0.270
(50,15)	0.764	0.022	0.092	0.769	0.379	0.030	0.033	0.381	−0.423	0.251	0.273	0.503	0.090	0.143	0.157	0.181
(20,50)	0.761	0.020	0.063	0.763	0.395	0.029	0.032	0.396	−0.299	0.222	0.314	0.434	0.348	0.160	0.211	0.407
(35,50)	0.757	0.015	0.058	0.759	0.386	0.020	0.021	0.387	−0.399	0.154	0.207	0.450	0.153	0.094	0.102	0.184
(50,50)	0.757	0.012	0.054	0.759	0.383	0.016	0.017	0.383	−0.418	0.127	0.171	0.452	0.099	0.077	0.079	0.127

Appendix C Data

Data are annual. All data are taken from OECD Economic Outlook (different years) except population data which are taken from OECD National Accounts Volume II Population and Employment (2009) and hours worked data which are taken from the Conference Board and Groningen Growth and Development Centre (2009). Data availability determines the sample period which is 1972-2007. The sample contains 15 countries: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Italy, Japan, the Netherlands, Spain, Sweden, the United Kingdom, and the United States. To calculate aggregate private consumption (C) we deflate private final consumption expenditures by the consumer price index.[14] To calculate the real interest rate (R) we subtract from the short-term nominal interest rate[15] the inflation rate which is calculated as the growth rate of the consumer price index. To calculate government consumption (G) we deflate government final consumption expenditures by the consumer price index. For aggregate hours worked (H) we use the series total hours worked as reported by the Conference Board. To calculate aggregate disposable labour income (Y) we first add the following three components. The first component is compensation of employees (a) which contains wages of the private sector as well as government wages and the social security contributions paid by private employers. The second component is the labour income of the self-employed (b) which we calculate as in Fiorito and Padrini (2001) by multiplying wages and salaries by the ratio of the number of self-employed to total employees. The third component is net social security transfers paid by the government (c), i.e. social security transfers paid by the government minus social security contributions received by the government. From (a)+(b)+(c) we then subtract taxes. To calculate taxes we follow Carey and Rabesona (2004) and make a distinction between countries where households cannot deduce their social security contributions from their tax base (Australia, Canada, United Kingdom, United States) and countries where households can deduce their social security contributions (all other countries). For the first group of countries the tax rate (d1) can be calculated as direct taxes on households divided by the sum of wages and salaries, property income received by households and total income of the self-employed. For the first group of countries the tax base (e1) is the sum of wages and salaries and labour income of the self-employed. Total

[14] For a few countries we use the deflator of private final consumption expenditures instead.
[15] For most countries we use the treasury bill rate. In some instances we use the money market rate or the discount rate.

taxes for the first group then equal (d1) x (e1). For the second group of countries the tax rate (d2) can be calculated as direct taxes on households divided by compensation of employees plus property income received by households plus total income of the self-employed minus social security contributions received by the government. For the second group of countries the tax base (e2) is compensation of employees plus labour income of the self-employed minus social security contributions received by the government. Total taxes for the second group then equal (d2) x (e2). Aggregate disposable labour income (Y) for the first group then equals (a)+(b)+(c)-(d1) x (e1) deflated by the consumer price index. For the second group it equals (a)+(b)+(c)-(d2) x (e2) deflated by the consumer price index. The variables C, G, H and Y are all divided by population to obtain per capita figures.