An overlapping-generations model for PolicyEngine reform scoring*

Vahid Ahmadi†

July 2026

Abstract

This paper documents the overlapping-generations (OLG) member of the PolicyEngine Macro model suite: the open-source OG-Core dynamic general-equilibrium framework of DeBacker and Evans (2024a), deployed today through its United Kingdom country calibration OG-UK, with the United States calibration OG-USA in scope. The model tracks 80 annual age cohorts of heterogeneous lifetime-ability households who choose labour supply and saving over the life cycle, a perfectly competitive production sector, and a government that taxes, transfers, spends, and issues debt, and it solves both the long-run stationary steady state and the year-by-year transition path between steady states. Its distinguishing feature within the suite is the microdata bridge of DeBacker et al. (2019): statutory reforms are specified as PolicyEngine tax–benefit parameters, run through the PolicyEngine UK microsimulation model over enhanced Family Resources Survey microdata, and summarised as smooth parametric effective- and marginal-tax-rate functions that the optimising households respond to — so the general-equilibrium model sees the true shape of a reform across the income distribution rather than a stylised wedge. I set out the model's equations, the tax-function estimation pipeline, the UK calibration and how selected targets compare with official ONS and OBR figures, and the `policyengine-macro` command-line integration (`pe-macro score --model og`). I am explicit about status: the wiring is experimental-grade, scoring is steady-state-only in the CLI today, a single solve takes upwards of seventeen minutes so the model is excluded from the hosted service, and — unlike the suite's OBR and SVAR members — there is no published ground truth to replicate: the model delivers structural counterfactuals, not validated forecasts.

Keywords: overlapping generations, dynamic general equilibrium, fiscal policy, microsimulation, tax functions, dynamic scoring

*Acknowledgements: I thank Max Ghenis for helpful discussions and suggestions. The OG-Core framework and its country calibrations are the work of the Policy Simulation Library community, in particular Jason DeBacker and Richard W. Evans; this paper documents PolicyEngine Macro's integration of that work and any errors in the description are mine. This paper was prepared with the assistance of Claude (Anthropic); all analysis and conclusions were reviewed by the author.

†Research Associate at PolicyEngine. Email: vahid@policyengine.org.

1 Introduction

PolicyEngine Macro is a suite of open-source macroeconomic models built around a single reform vocabulary: a tax–benefit reform is a flat dictionary of PolicyEngine parameter changes, and every member of the suite consumes that same object through its own declared contract. The suite's other members are anchored to published figures — an emulator of the Office for Budget Responsibility's (OBR) macroeconometric model, a Bank-of-England-style structural VAR, and the PolicyEngine static microsimulation itself. This paper documents the member that answers a different kind of question: the *psl-og* model, a dynamic overlapping-generations (OLG) general-equilibrium model built on the Policy Simulation Library's OG-Core framework (DeBacker and Evans, 2024a), wired into the suite today through its United Kingdom country calibration OG-UK (DeBacker et al., 2025), with the United States calibration OG-USA (DeBacker and Evans, 2024b) in scope.

The model is a large-scale descendant of Auerbach and Kotlikoff (1987). It tracks eighty annual age cohorts — households aged 20 to 99, entering economic life at 20 and facing age-specific mortality — each populated by several lifetime-ability types. Every cohort chooses labour supply and saving over its remaining lifetime, trading off consumption against an elliptical disutility of work and a warm-glow bequest motive. Perfectly competitive firms rent the resulting capital and labour; the government taxes, transfers, spends, and issues debt, with a fiscal rule that stabilises the long-run debt-to-GDP ratio. Two solution objects are on offer: the *stationary steady state* ("where does the economy settle under this policy?") and the *transition path* (a year-by-year solve, over a multi-decade horizon, of how it gets there).

Within the suite, the OLG member's role is long-run structural analysis. The OBR emulator and the SVAR speak to demand-side dynamics over a forecast window; the static microsimulation speaks to first-round revenue and distribution. The OLG model is the only member in which a tax reform changes prices — wages, interest rates — through households' labour-supply and saving responses, so it is the member that can address questions like the long-run output cost of a higher basic rate of income tax, the crowding-out effects of debt, or the incidence of corporation tax across generations. Its distinguishing input channel is the microdata bridge of DeBacker et al. (2019): reforms are not stylised wedges but actual statute, run through the PolicyEngine UK microsimulation model (Woodruff et al., 2025) over enhanced Family Resources Survey (FRS) microdata and compressed into smooth parametric effective- and marginal-tax-rate functions that the optimising households face.

Two properties of this member must be stated at the outset, because they shape everything downstream. First, it is *local-only*: a single steady-state solve — microdata calibration plus the general-equilibrium fixed point — has been measured at over seventeen minutes on a laptop, and a full transition path takes hours, so the model is deliberately excluded from the hosted PolicyEngine Macro service and runs only on the user's own machine through the `pe-macro` command-line interface. Second, unlike the suite's OBR and SVAR members, it is a *calibrated structural model with no published ground truth to replicate.* Its outputs are structural counterfactuals — internally consistent answers conditional on the model's assumptions — not validated forecasts. This paper is accordingly written as documentation with explicit caveats (Section 8), not as a validation exercise.

The remainder of the paper proceeds as follows. Section 2 places the model in the OLG literature and cites the framework's own documentation. Section 3 sets out the model's equations: the household problem, the firm problem, the government budget and fiscal rules, market clearing, and the steady-state and transition-path solution algorithms. Section 4 details the tax-function estimation pipeline from PolicyEngine microdata — the PolicyEngine-specific contribution. Section 5 describes the software architecture and the `policyengine-macro` integration. Section 6 documents the UK calibration and compares selected targets with official statistics. Section 8 states the model's limitations honestly, and Section 9 concludes. Three appendices supply a parameter glossary, the OG-Core API surface the integration touches, and an exact pinned-version reproduction recipe. The figures throughout are generated from the calibration's cheap-to-load inputs — the UN demographic series, the OG-Core ability matrix, the fitted utility and tax functional forms — and require no equilibrium solve to reproduce (`figures/make_figures.py`).

2 Related literature and model provenance

The psl-og member sits at the intersection of four literatures: the overlapping-generations tradition in public finance, the micro–macro linkage literature that connects microsimulation to structural models, UK applications of life-cycle and OLG modelling, and the open-source policy-modelling ecosystem from which the software itself comes.

2.1 The OLG lineage

The intellectual ancestry begins with Samuelson (1958), whose consumption-loan economy established that an economy of finitely lived, overlapping cohorts can support equilibria — including dynamically inefficient ones — that have no counterpart in infinitely-lived-agent models, and Diamond (1965), who embedded the structure in a neoclassical growth model and showed how national debt shifts consumption across generations, the comparative-static heart of every fiscal-policy OLG exercise since. The theory of lifetime uncertainty and annuities that underlies the model's mortality-weighted discounting and accidental bequests goes back to Yaari (1965).

The computational turn came with Auerbach and Kotlikoff (1987): a 55-period, perfect-foresight, general-equilibrium life-cycle model solved by what is now called time-path iteration, applied to tax reform, social security, and generational incidence. Essentially every quantitative claim in that book's genre — long-run output effects of moving from income to consumption taxation, the intergenerational burden of debt — is the kind of claim OG-Core produces today, and the solution method is a direct descendant (Section 3.7.2; Evans and Phillips, 2014). The ambitious-heterogeneity successors extended the framework in the directions the modern model inherits: İmrohoroğlu et al. (1995) added idiosyncratic earnings risk and incomplete markets to social-security analysis; Nishiyama and Smetters (2007) combined idiosyncratic wage shocks with a rich social-security institution to weigh privatisation's efficiency gains against lost risk sharing; Conesa et al. (2009) used a lifecycle economy with ability heterogeneity to revisit optimal capital taxation; and De Nardi (2004) established the warm-glow bequest specification with heterogeneous intensity that OG-Core's χ_j^b term follows. On the labour-supply margin, the model's calibrated Frisch elasticity engages the long micro–macro elasticity debate surveyed by Blundell and MaCurdy (1999) and Blundell et al. (2007) and sharpened by Peterman (2016), who shows how life-cycle aggregation reconciles small micro estimates with the larger elasticities macro calibrations need. Large applied OLG-CGE systems in the consulting-and-scoring tradition — e.g. the Zodrow–Diamond model used for US corporate-tax analysis (Zodrow and Diamond, 2013) and the multi-region Fehr–Kotlikoff-style demographic models (Kotlikoff et al., 2007) — are the closed-source cousins of what OG-Core does in the open.

2.2 Micro–macro linkage

A second strand connects tax microsimulation — the institution-rich, behaviour-free calculators that ministries and think tanks run on survey microdata — with structural macro models that have behaviour but stylised institutions. The methodological bridge this paper's pipeline implements is DeBacker et al. (2019), who propose estimating parametric effective- and marginal-tax-rate functions, jointly in labour and capital income, from the output of a microsimulation model, and embedding those functions in the DGE household problem. The functional-form menu draws on the empirical tax-function literature: the three-parameter total-income schedule of Gouveia and Strauss (1994), long used to summarise US federal taxes in calibrated macro models, and the log-linear progressivity form of Heathcote et al. (2017), the workhorse of the analytical optimal-progressivity literature, alongside DeBacker–Evans–Phillips's own richer bivariate specification (Section 4.2). PolicyEngine Macro's contribution is not to this theory but to the pipeline: replacing the microsimulation input with PolicyEngine's open-source UK model over enhanced FRS microdata, and exposing the whole loop — reform in, steady-state general-equilibrium impact out — behind the suite's common scoring interface.

2.3 UK applications

The nearest UK relative is the Office for Budget Responsibility's own overlapping-generations model, documented in OBR Working Paper 22 (Office for Budget Responsibility, 2025b): an OLG economy with idiosyncratic earnings risk (and perfect foresight over aggregates) built for long-run fiscal-sustainability analysis, sharing this model's broad architecture (annual cohorts, life-cycle optimisation, a government block closed on a debt rule) while differing in calibration detail and in institutional embedding — the OBR model is anchored to the official long-term fiscal projections, where OG-UK is anchored to PolicyEngine microdata. Because the two models answer overlapping questions from independent codebases, side-by-side comparison against the OBR model is the closest available substitute for validation of the psl-og member (Section 8). The UK also has a deep empirical life-cycle literature that disciplines the model's behavioural margins: the IFS tradition of estimating life-cycle labour supply and human capital on UK panel data — e.g. Blundell et al. (2016), whose structural female labour-supply model quantifies exactly the participation and hours responses to taper rates and in-work benefits that the estimated MTR functions of Section 4 summarise — and the Frisch-elasticity evidence surveyed in Blundell et al. (2007). On the demographic side, the model consumes the same UN and ONS projections (United Nations, Department of Economic and Social Affairs, Population Division, 2024; Office for National Statistics, 2024) that drive the OBR's Fiscal Risks and Sustainability analysis, so its ageing-driven long-run dynamics are commensurable with official work.

2.4 The open-source model ecosystem

OG-Core is the open-source embodiment of the Auerbach–Kotlikoff tradition, maintained by the Policy Simulation Library (PSL) community — a federation of open, peer-reviewed policy models with common packaging and documentation standards (Policy Simulation Library, 2024a). Its theory, calibration interface, solution methods, and Python API are documented in DeBacker and Evans (2024a); Evans and DeBacker (2024) provide a textbook treatment of building and solving OLG models for policy analysis covering most of the structures in OG-Core, and Evans and Phillips (2014) analyse the time-path-iteration solution method. The framework separates *theory* (OG-Core, a country-agnostic system of equations and solvers) from *calibration* (country packages supplying demographics, tax detail, and macro parameters). OG-USA (DeBacker and Evans, 2024b) is the United States calibration, used with the PSL Tax-Calculator microsimulation (Policy Simulation Library, 2024b) for dynamic scoring of US federal tax reform — the original instance of the microsimulation-to-DGE pipeline this paper describes; sister calibrations exist for several other countries. OG-UK (DeBacker et al., 2025) is the United Kingdom calibration, originally built on OpenFisca-UK and now on PolicyEngine UK (Woodruff et al., 2025), sourcing calibration targets from the ONS, the OBR, the Bank of England, and GOV.UK. PolicyEngine itself extends the same open-source ethos to the microsimulation layer: versioned statute-as-code, public microdata enhancement, and reproducible reform APIs, of which the psl-og member is the general-equilibrium consumer.

3 The model

This section sets out the blocks of OG-Core following DeBacker and Evans (2024a), in the configuration used by the PolicyEngine Macro integration: a single production sector, and households differentiated by age and lifetime ability. Relative to the compact statement in earlier drafts, this section derives the household first-order conditions step by step, motivates the elliptical disutility of labour against the constant-Frisch-elasticity (CFE) alternative, derives the firm conditions under the corporate tax wedge, states the full government block including every fiscal closure rule the framework supports, and describes the demographic machinery and the two solution algorithms with their convergence criteria.

Time is annual. Households live at most $E+S$ periods: E economically inactive childhood years (ages 0–19) and $S=80$ active years (model ages $s=E+1,\dots,E+S$, real ages 20–99). Each cohort contains J lifetime-ability types indexed by j, with population shares λ_j ($\sum_j \lambda_j = 1$) and deterministic age–ability effective-labour profiles $e_{j,s}$. Population $\omega_{s,t}$ evolves by exogenous fertility, mortality ρ_s, and immigration (Section 3.1); labour-augmenting productivity grows at rate g_y, and all equations below are expressed before stationarisation (Section 3.6).

3.1 Demographics

The demographic block is exogenous to the equilibrium but is the engine of the model's long-run dynamics, so we state it in full. Let $\omega_{s,t}$ denote the number of individuals of age $s \in \{1,\dots,E{+}S\}$ alive at time t, f_s the age-specific fertility rate (births per person of age s per year), ρ_s the age-specific mortality hazard (probability of dying between ages s and $s+1$, with $\rho_{E+S}=1$ imposed so that no one outlives the model), and i_s the age-specific net immigration rate expressed as a fraction of the existing age-s population. The population laws of motion are

$$\omega_{1,t+1} \;=\; (1-\rho_0)\sum_{s=1}^{E+S} f_s\,\omega_{s,t} \;+\; i_1\,\omega_{1,t}, \tag{1}$$

$$\omega_{s+1,t+1} \;=\; (1-\rho_s)\,\omega_{s,t} \;+\; i_{s+1}\,\omega_{s+1,t}, \qquad 1 \le s \le E+S-1, \tag{2}$$

where ρ_0 is infant mortality. Equation (1) says next year's newborns are this year's births that survive infancy plus net child migration; equation (2) says each cohort ages one year, shrinks by mortality, and is augmented by net migration into its age cell.

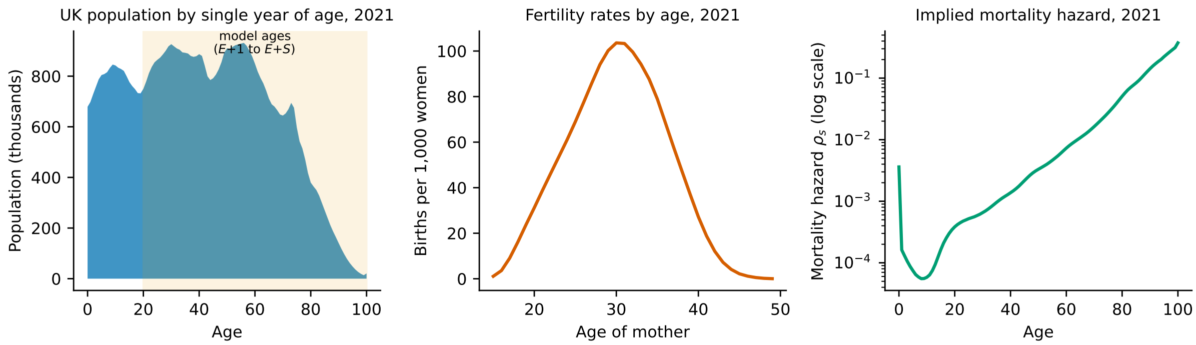

Figure 1: Demographic inputs to OG-UK, from the UN World Population Prospects single-year-of-age series for the United Kingdom shipped with the `oguk` package (United Nations, Department of Economic and Social Affairs, Population Division, 2024). Left: population by single year of age (2021), with the model's economically active ages 20–99 shaded. Centre: fertility rates by age of mother. Right: the implied mortality hazard ρ_s (deaths divided by population, log scale), which enters both the households' effective discount factor and the warm-glow bequest weight.

Data and the immigration residual. OG-UK populates $\{f_s, \rho_s, \omega_{s,t}\}$ from the United Nations World Population Prospects single-year-of-age series for the United Kingdom (United Nations, Department of Economic and Social Affairs, Population Division, 2024), retrieved through the UN Data Portal API with a local CSV fallback shipped in the package (population, fertility, deaths, and infant mortality by single year of age; Figure 1 plots the three inputs). Immigration rates i_s are not taken from a migration series directly; they are computed as the *residual* that reconciles two successive observed population vectors given observed fertility and mortality — solving (1)–(2) for i_s given $(\omega_{\cdot,t}, \omega_{\cdot,t+1}, f_s, \rho_s)$ and averaging over recent year pairs. This guarantees the model's near-term population path tracks the official projection exactly, at the cost of attributing any residual measurement noise to migration.

Transition to a stationary population. A steady state requires a time-invariant population age distribution. The framework therefore constructs the population path in three phases: (i) for the first $\approx$50 transition periods, the raw laws of motion (1)–(2) run forward from the current observed age distribution; (ii) the age distribution is then linearly interpolated towards the stationary distribution implied by fixed (f_s, ρ_s, i_s) — the normalised Perron eigenvector of the linear map defined by (1)–(2), whose dominant eigenvalue is one plus the stationary population growth rate $\bar{g}_n$; and (iii) from period T onward the distribution is exactly stationary. The economically active shares used by the equilibrium are the normalised $\hat{\omega}_{s,t} = \omega_{s,t} / \sum_{u=E+1}^{E+S} \omega_{u,t}$ over model ages, and the period-t population growth rate $g_{n,t}$ enters the stationarised resource constraint and the capital-market clearing condition below. The ageing embedded in the UK projection — a rising old-age share for the next several decades — is thus an input the steady state inherits, not an assumption bolted on.

3.2 Households

A household of type j and age s at time t has period utility that is additively separable in consumption $c_{j,s,t}$, labour $n_{j,s,t}$, and bequests-to-be-left $b_{j,s+1,t+1}$:

$$u(c_{j,s,t}, n_{j,s,t}, b_{j,s+1,t+1}) \;=\; \frac{c_{j,s,t}^{1-\sigma} - 1}{1-\sigma} \;+\; e^{g_y t(1-\sigma)}\, \chi_s^n\, v(n_{j,s,t}) \;+\; \chi_j^b\, \rho_s\, \frac{b_{j,s+1,t+1}^{1-\sigma} - 1}{1-\sigma}, \tag{3}$$

where σ is the coefficient of relative risk aversion, χ_s^n is an age-specific disutility-of-labour scale, and the warm-glow bequest term is weighted by the mortality hazard ρ_s and an ability-specific intensity χ_j^b — a "joy of giving" motive in the tradition of De Nardi (2004), scaled by the probability that the bequest is actually left this period. The factor $e^{g_y t(1-\sigma)}$ on the labour term keeps the consumption–leisure trade-off stable along the balanced-growth path: consumption grows at rate g_y while the time endowment does not, so without this scaling the marginal-utility ratio would trend and hours would drift.

3.2.1 Elliptical disutility of labour, and why not CFE

The workhorse specification in the quantitative literature is the constant-Frisch-elasticity (CFE) form

$$v_{\mathrm{CFE}}(n) \;=\; -\frac{n^{1+\frac{1}{\theta}}}{1+\frac{1}{\theta}}, \qquad v'_{\mathrm{CFE}}(n) \;=\; n^{1/\theta}, \tag{4}$$

whose Frisch elasticity of labour supply is the constant θ at every interior point. Its computational defect in a life-cycle model is at the boundaries: $v'_{\mathrm{CFE}}(0) = 0$, so nothing in the preferences stops the optimiser from choosing $n \le 0$ for cohorts whose after-tax wage is low (the very old, the low-ability types), and nothing stops n exceeding the time endowment $\tilde{l}$ for high-wage cohorts. Handling this properly requires Kuhn–Tucker complementary-slackness conditions at both bounds for every one of the $S \times J$ labour choices — occasionally binding constraints that turn a smooth root-finding problem into a nonsmooth one and are a well-documented source of solver failure.

OG-Core instead uses the *elliptical* disutility of Evans and Phillips (2018): take the upper-right quadrant of an ellipse in $(n, -v)$ space with horizontal radius $\tilde{l}$,

$$v(n) \;=\; -b\left[1 - \left(\frac{n}{\tilde{l}}\right)^{\upsilon}\right]^{1/\upsilon}, \qquad b, \upsilon > 0, \tag{5}$$

with marginal disutility

$$v'(n) \;=\; \frac{b}{\tilde{l}}\left(\frac{n}{\tilde{l}}\right)^{\upsilon-1}\left[1 - \left(\frac{n}{\tilde{l}}\right)^{\upsilon}\right]^{\frac{1}{\upsilon}-1}. \tag{6}$$

For $\upsilon > 1$, expression (6) tends to 0^+ as $n \to 0$ *slower than any interior CFE schedule near the top* and, crucially, tends to $+\infty$ as $n \to \tilde{l}$: the Inada-type behaviour at the endowment bound means neither constraint ever binds at an optimum, interior first-order conditions are sufficient everywhere, and the Kuhn–Tucker apparatus disappears. The parameters (b, υ) are not free: OG-Core's `elliptical_u_est` module fits them by minimising the squared distance between (6) and the CFE marginal disutility $n^{1/\theta}$ over an interior grid of n values, given the calibrated Frisch elasticity θ. The fitted ellipse thus *mimics* CFE behaviour where the data live — interior hours — while repairing its boundary behaviour (Figure 2). With the OG-Core default Frisch elasticity the fit delivers $b \approx 0.573$ and $\upsilon \approx 2.856$.

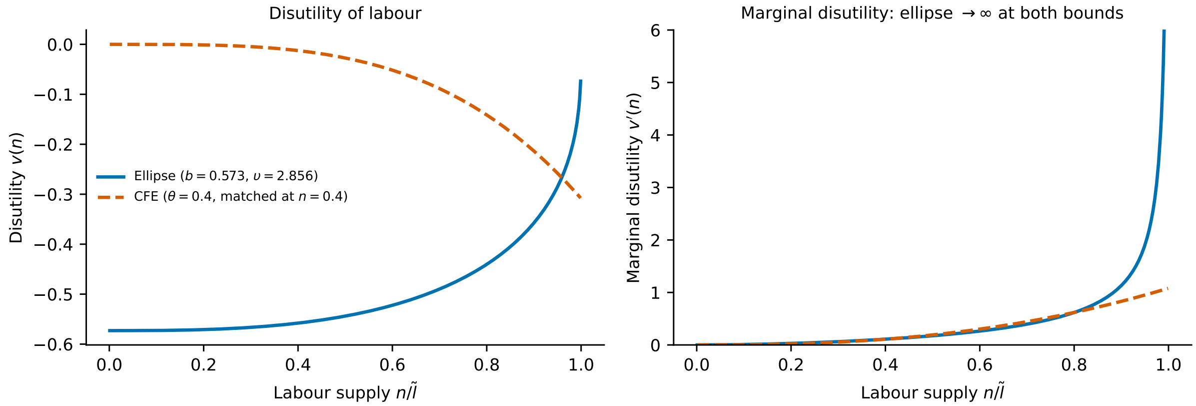

Figure 2: The elliptical disutility of labour (5) against a CFE specification calibrated to the same Frisch elasticity (scale matched at $n = 0.4$), computed from the fitted OG-Core default parameters (b, υ). Left: disutility levels. Right: marginal disutility — the ellipse tracks CFE over interior hours but its marginal disutility diverges to infinity as $n \to \tilde{l}$ (and its curvature keeps $n > 0$), so labour-supply bounds never bind and Kuhn–Tucker conditions are unnecessary (Evans and Phillips, 2018).

3.2.2 Budget constraint

The period budget constraint, in the single-good configuration used here (suppressing the multi-good consumption-tax detail of the general framework), is

$$(1+\tau_t^c)\, c_{j,s,t} + b_{j,s+1,t+1} \;=\; (1+r_{p,t})\, b_{j,s,t} + w_t\, e_{j,s}\, n_{j,s,t} + bq_{j,s,t} + tr_{j,s,t} + pension_{j,s,t} - tax_{j,s,t}, \quad (7)$$

where $b_{j,s,t}$ is beginning-of-period wealth, w_t is the wage per effective labour unit, $bq_{j,s,t}$ are received bequests, $tr_{j,s,t}$ are government transfers, and $tax_{j,s,t} = \tau^{etr}(x_{j,s,t}, y_{j,s,t})\,(x_{j,s,t} + y_{j,s,t})$ is the net income-tax liability generated by the effective-tax-rate function of Section 4, evaluated at labour income $x_{j,s,t} \equiv w_t e_{j,s} n_{j,s,t}$ and capital income $y_{j,s,t} \equiv r_{p,t} b_{j,s,t}$. Two prices deserve comment. The household's return $r_{p,t}$ is the return on a *portfolio*: households hold both private capital and government bonds in proportion to their aggregate supplies, so

$$r_{p,t} \;=\; \frac{r_t\, K_t + r_{gov,t}\, D_t}{K_t + D_t}, \quad (8)$$

where r_t is the after-corporate-tax return on capital from the firm block and $r_{gov,t}$ is the government borrowing rate of Section 3.4. Accidental bequests — the wealth of those who die, $\sum_s \rho_s\, \omega_{s,t}\, \lambda_j\, b_{j,s+1,t+1}$ — are redistributed lump-sum within the same ability type j (the deployed configuration; the framework also supports redistribution across types by a fixed matrix), so $bq_{j,s,t}$ is the type-j per-capita share of aggregate type-j bequests $BQ_{j,t}$.

3.2.3 The household problem and first-order conditions, step by step

A household entering economic life at time t chooses $\{c_{j,s,u}, n_{j,s,u}, b_{j,s+1,u+1}\}$ over its remaining lifetime to maximise

$$\sum_{u=0}^{S-1} \beta^u \left[\prod_{v=0}^{u-1} (1 - \rho_{E+1+v})\right] u\big(c_{j,E+1+u,t+u},\; n_{j,E+1+u,t+u},\; b_{j,E+2+u,t+u+1}\big), \quad (9)$$

subject to the sequence of budget constraints (7), where β is the pure discount factor and the product of survival probabilities is the probability of being alive to enjoy period-u utility. There is no aggregate or idiosyncratic risk beyond mortality, and annuity markets are absent (hence accidental bequests). Substitute the budget constraint into (9) to eliminate $c_{j,s,t}$, or equivalently attach multiplier $\mu_{j,s,t}$ to each period's constraint; we take the Lagrangian route for transparency.

Step 1: marginal utility of wealth. Differentiating the Lagrangian with respect to $c_{j,s,t}$ gives

$$(c_{j,s,t})^{-\sigma} \;=\; (1+\tau_t^c)\,\mu_{j,s,t}, \tag{10}$$

so the multiplier is the marginal utility of consumption deflated by the consumption-tax wedge.

Step 2: labour supply. Differentiating with respect to $n_{j,s,t}$: labour raises resources by the after-tax marginal wage. Because *tax* depends on labour income both through the ETR level and through the ETR's own slope, the relevant derivative is the *marginal* tax rate on labour income,

$$\tau_{j,s,t}^{mtrx} \;\equiv\; \frac{\partial\, tax_{j,s,t}}{\partial x_{j,s,t}} \;=\; \tau^{etr} + (x+y)\,\frac{\partial \tau^{etr}}{\partial x}, \tag{11}$$

which the pipeline of Section 4 estimates directly rather than differentiating the fitted ETR (except under Gouveia–Strauss, where the derivative is analytic). The first-order condition is

$$w_t e_{j,s}\big(1-\tau_{j,s,t}^{mtrx}\big)\,\mu_{j,s,t} \;=\; e^{g_y t(1-\sigma)}\,\chi_s^n\,v'(n_{j,s,t}), \tag{12}$$

and substituting (10) for the multiplier gives the labour-supply Euler equation

$$\frac{w_t e_{j,s}\big(1-\tau_{j,s,t}^{mtrx}\big)}{1+\tau_t^c}\,(c_{j,s,t})^{-\sigma} \;=\; e^{g_y t(1-\sigma)}\,\chi_s^n\,\frac{b}{\tilde{l}}\left(\frac{n_{j,s,t}}{\tilde{l}}\right)^{\upsilon-1}\left[1-\left(\frac{n_{j,s,t}}{\tilde{l}}\right)^{\upsilon}\right]^{\frac{1}{\upsilon}-1}. \tag{13}$$

The left side is the utility value of the after-tax, after-consumption-tax real wage; the right side is the elliptical marginal disutility (6). This is where the estimated τ^{mtrx} function bites: a reform that raises marginal rates over some income range rotates this condition for exactly the households whose (x,y) fall in that range.

Step 3: saving. Differentiating with respect to $b_{j,s+1,t+1}$, which appears in three places — today's budget constraint (cost: one unit of resources), tomorrow's budget constraint (benefit: principal plus after-tax return, weighted by survival), and today's warm-glow term (benefit: marginal bequest utility, weighted by death) — gives

$$\mu_{j,s,t} \;=\; \beta(1-\rho_s)\Big[1+r_{p,t+1}\big(1-\tau_{j,s+1,t+1}^{mtry}\big)\Big]\,\mu_{j,s+1,t+1} \;+\; \chi_j^b\,\rho_s\,(b_{j,s+1,t+1})^{-\sigma}, \tag{14}$$

where τ^{mtry} is the marginal rate on capital income defined analogously to (11). Substituting (10) (and assuming a constant consumption-tax rate across the two periods for readability) yields the consumption–saving Euler equation for $s < E+S$:

$$(c_{j,s,t})^{-\sigma} \;=\; \beta(1-\rho_s)\Big[1+r_{p,t+1}\big(1-\tau_{j,s+1,t+1}^{mtry}\big)\Big]\,(c_{j,s+1,t+1})^{-\sigma} \;+\; (1+\tau_t^c)\,\chi_j^b\,\rho_s\,(b_{j,s+1,t+1})^{-\sigma}. \tag{15}$$

The bequest term acts as a mortality-weighted floor on the return to saving: even when survival-weighted future consumption is worth little (old age, high ρ_s), assets retain marginal value because they may be bequeathed. This is what keeps end-of-life asset holdings positive and generates the aggregate bequest flow BQ.

Step 4: terminal condition. In the final period of life $s = E + S$, $\rho_{E+S} = 1$: there is no survival continuation, and the choice of terminal assets $b_{j,E+S+1,t+1}$ trades consumption directly against the warm glow,

$$(c_{j,E+S,t})^{-\sigma} \;=\; (1+\tau_t^c)\,\chi_j^b\,(b_{j,E+S+1,t+1})^{-\sigma}, \tag{16}$$

which is (15) evaluated at $\rho_s = 1$. Equations (13), (15) and (16) — $2S$ nonlinear equations per ability type, given prices and aggregates — fully characterise household behaviour, and they show precisely where the microdata-estimated tax functions enter: reforms move τ^{etr}, τ^{mtrx} and τ^{mtry} across the joint income distribution, and the general equilibrium reprices w_t and r_t in response.

3.3 Firms

A unit measure of perfectly competitive firms produces the single output with constant-elasticity-of-substitution technology in private capital K_t and effective labour L_t (the deployed configuration switches off public-capital augmentation):

$$Y_t \;=\; F(K_t, L_t) \;=\; Z_t\Big[\gamma^{1/\varepsilon} K_t^{\frac{\varepsilon-1}{\varepsilon}} + (1-\gamma)^{1/\varepsilon}\big(e^{g_y t} L_t\big)^{\frac{\varepsilon-1}{\varepsilon}}\Big]^{\frac{\varepsilon}{\varepsilon-1}}, \tag{17}$$

where Z_t is total factor productivity, γ is the capital share parameter, and ε is the elasticity of substitution. In the Cobb–Douglas limit $\varepsilon = 1$, (17) reduces to $Y_t = Z_t K_t^{\gamma}\big(e^{g_y t} L_t\big)^{1-\gamma}$.

Firms face a corporate income tax at rate τ_t^b on revenue net of labour costs, with a deduction for depreciation at the *tax* depreciation rate δ_t^τ (which may differ from economic depreciation δ — capital allowances such as full expensing correspond to $\delta^\tau > \delta$). Firms rent capital at $r_t + \delta$ (the household's required return plus replacement of depreciation) and choose K_t, L_t each period to maximise after-tax profits

$$PR_t \;=\; (1-\tau_t^b)\big(Y_t - w_t L_t\big) - \big(r_t + \delta\big)K_t + \tau_t^b\,\delta_t^\tau K_t. \tag{18}$$

The problem is static because capital is rented, not owned. Differentiating with respect to L_t:

$$(1-\tau_t^b)\Big(\frac{\partial F}{\partial L_t} - w_t\Big) = 0 \quad\Longrightarrow\quad w_t \;=\; \frac{\partial F}{\partial L_t} \;=\; e^{g_y t}\, Z_t^{\frac{\varepsilon-1}{\varepsilon}}\Big[(1-\gamma)\,\frac{Y_t}{e^{g_y t} L_t}\Big]^{1/\varepsilon}. \tag{19}$$

Because wages are deductible, the corporate tax does not distort the labour-demand margin: the wage equals the marginal product of labour exactly. Differentiating with respect to K_t:

$$(1-\tau_t^b)\,\frac{\partial F}{\partial K_t} - (r_t + \delta) + \tau_t^b\,\delta_t^\tau = 0 \quad\Longrightarrow\quad r_t \;=\; (1-\tau_t^b)\,\Big[\gamma\,\frac{Y_t}{K_t}\Big]^{1/\varepsilon} Z_t^{\frac{\varepsilon-1}{\varepsilon}} - \delta + \tau_t^b\,\delta_t^\tau. \tag{20}$$

The corporate tax drives a wedge between the marginal product of capital and the return households receive: a higher τ^b lowers r_t for given K_t, households accumulate less capital, and in the new equilibrium the pre-tax marginal product is higher, output lower, and — through (19) — wages lower. This is the channel by which corporation-tax incidence falls partly on labour in the long run, a question the OLG member is specifically equipped to quantify. When $\delta^\tau = \delta$ the depreciation deduction exactly compensates for economic depreciation of the tax base; full expensing (δ^τ effectively covering the purchase) pushes the system towards a cash-flow tax that leaves the marginal investment undistorted. The corporation-tax rate τ^b is a structural parameter of the macro model rather than a PolicyEngine parameter; in the integration it is shocked directly through parameter overrides (Section 5).

3.4 Government

The government levies household income taxes through the estimated tax functions, consumption taxes at rate τ_t^c, wealth taxes if configured (off in the deployed configuration), and the corporate tax; it spends on public consumption G_t, non-pension transfers TR_t, public infrastructure investment $I_{g,t}$ (off in the deployed configuration), and pensions; and it issues one-period debt D_t.

Revenue. Aggregate revenue sums the household income-tax liabilities generated by the fitted ETR functions, consumption-tax receipts, and business-tax receipts:

$$Rev_t \;=\; \underbrace{\sum_{s,j} \omega_{s,t}\lambda_j \, \tau^{etr}(x_{j,s,t}, y_{j,s,t})\,(x_{j,s,t}+y_{j,s,t})}_{\text{income taxes}} \;+\; \underbrace{\tau_t^c\, C_t}_{\text{consumption taxes}} \;+\; \underbrace{\tau_t^b\big(Y_t - w_t L_t\big) - \tau_t^b\,\delta_t^\tau K_t}_{\text{business taxes}}. \tag{21}$$

Debt and the government interest rate. Debt is one-period and rolls over at rate $r_{gov,t}$, which the framework allows to differ from the marginal product-based r_t through an affine wedge,

$$r_{gov,t} \;=\; \tau_{d,t}\, r_t + \mu_{d,t}, \tag{22}$$

with scale τ_d and level μ_d calibrated so that the model's effective interest rate on government debt matches the observed gilt-implied rate rather than the corporate return. The flow budget constraint is

$$D_{t+1} + Rev_t \;=\; (1 + r_{gov,t})\, D_t + G_t + TR_t + Pensions_t + I_{g,t}. \tag{23}$$

Baseline spending rules. Along the baseline, non-pension transfers and public consumption follow fixed GDP shares, $TR_t = \alpha_{tr} Y_t$ and $G_t = \alpha_g Y_t$ (and $I_{g,t} = \alpha_I Y_t$ when public investment is on); pensions follow the configured pension rule (the deployed configuration uses OG-Core's notional-defined-contribution/flat options set to approximate the UK state pension as an age-conditioned transfer).

Fiscal closure rules. Equation (23) does not by itself pin down a stationary debt path: with spending shares fixed and revenue determined by behaviour, debt would generically follow an explosive or degenerate trajectory. Some fiscal instrument must eventually adjust — the model must be *closed*. OG-Core implements closure with a two-phase rule governed by parameters $(T_{G1}, T_{G2}, \alpha_D, \rho_G)$: before period T_{G1} the baseline rules run untouched (deficits are whatever they are, capturing near-term announced policy); between T_{G1} and T_{G2} the chosen instrument adjusts partially each period, e.g. for a spending closure

$$G_t \;=\; \rho_G\, \alpha_g Y_t \;+\; (1-\rho_G)\Big[(1+g_{n,t})e^{g_y}\,\alpha_D Y_t - (1+r_{gov,t})D_t + Rev_t - TR_t - Pensions_t\Big], \tag{24}$$

so that with smoothing weight $\rho_G \in [0, 1)$ the debt ratio glides towards the long-run target α_D; after T_{G2} debt is held exactly at $D_{t+1} = (1 + g_{n,t})e^{g_y}\alpha_D Y_t$ and the instrument absorbs the residual one-for-one. The framework supports the same structure with any of the following as closing instrument: public consumption G, non-pension transfers TR, both in fixed proportion, or (as a budget-balance option) period-by-period balance with debt fixed. The choice matters economically: closing on G (pure waste in the utility function) isolates the tax reform's incentive effects; closing on TR returns resources to households and mutes wealth effects. The deployed configuration closes on G, with UK-configured phase parameters that depart deliberately from the OG-Core defaults ($T_{G1} = 20$, $T_{G2} = 256$, $\rho_G = 0.1$): the deployed JSON sets $T_{G1} = 4$, $T_{G2} = 20$ and $\rho_G = 0.6$, so consolidation starts earlier, completes sooner, and adjusts faster. In the steady state, debt is exactly $\bar{D} = \alpha_D \bar{Y}$ and (23) determines the residual spending level consistent with it. Every reform score therefore embeds the closure rule: a tax cut that loses revenue is eventually paid for by the closing instrument, and the reported long-run effects are net of that financing.

3.5 Market clearing and equilibrium

Equilibrium requires three markets to clear. Labour supply aggregates to

$$L_t \;=\; \sum_{s=E+1}^{E+S} \sum_{j=1}^{J} \omega_{s,t}\, \lambda_j\, e_{j,s}\, n_{j,s,t}, \tag{25}$$

household savings fund private capital and government debt (in the closed-economy configuration; the small-open-economy features allow foreign holdings of a fraction of debt at an exogenous world rate r^*),

$$K_t + D_t \;=\; \sum_{s=E+2}^{E+S+1} \sum_{j=1}^{J} \omega_{s-1,t-1}\, \lambda_j\, b_{j,s,t}, \tag{26}$$

and the goods market clears by Walras's law:

$$Y_t \;=\; C_t + I_t + G_t, \qquad I_t = K_{t+1} - (1-\delta)K_t. \tag{27}$$

A *steady-state equilibrium* is a stationary allocation and price pair $(\bar{r}, \bar{w})$ such that households optimise, firms optimise, the government satisfies its closure rule, and all markets clear with time-invariant stationarised quantities. A *non-steady-state equilibrium* is a full time path of allocations and prices from arbitrary initial conditions converging to that steady state.

3.6 Stationarisation

The raw system trends for two reasons: labour-augmenting productivity growth g_y and population growth $g_{n,t}$. Per-capita per-effective-unit variables are stationarised by deflating: individual quantities $\hat{c} = c\, e^{-g_y t}$, $\hat{b} = b\, e^{-g_y t}$; aggregates additionally by the population index N_t, $\hat{Y}_t = Y_t e^{-g_y t}/N_t$ and likewise $\hat{K}, \hat{C}, \hat{D}$; labour only by population, $\hat{L}_t = L_t/N_t$ (hours per person do not grow); and prices are already stationary except that the wage deflates to $\hat{w}_t = w_t e^{-g_y t}$. Substituting transforms, e.g., the Euler equation (15) into its stationary equivalent with β replaced by the growth-adjusted discount factor $\beta\, e^{-\sigma g_y}$, and the capital law of motion into $(1 + g_{n,t+1}) e^{g_y} \hat{K}_{t+1} = (1-\delta)\hat{K}_t + \hat{I}_t$. All solving happens in hatted variables; the integration's £bn mapping (Section 5) reverses the normalisation against ONS levels.

3.7 Solution algorithms

3.7.1 Steady state

The steady state is computed as a fixed point in a small vector of aggregates. Collect the outer unknowns $\Theta = (\bar{r},\ \bar{w},\ \overline{BQ}_j,\ \overline{TR},\ \text{factor})$ — the interest rate, wage, aggregate bequests by type, aggregate transfers, and the model-to-data income scaling *factor* used to evaluate the tax functions (which are estimated in pounds) at model-unit incomes. The algorithm is:

1. **Guess** $\Theta^{(0)}$.
2. **Inner solve.** Given $\Theta^{(i)}$, factor prices are known; solve each ability type's stationary lifetime problem — the $2S$ equations (13), (15), (16) in $(\bar{n}_{j,s}, \bar{b}_{j,s+1})_{s=E+1}^{E+S}$ — with a Newton-type root finder (`scipy.optimize.root`), warm-started from the previous outer iteration's policies.
3. **Aggregate.** Sum the resulting policies over (j, s) with stationary weights $\hat{\omega}_s \lambda_j$ to get implied $\hat{K}, \hat{L}, \widehat{BQ}_j$; evaluate firm conditions (19)–(20) at $(\hat{K}, \hat{L})$ to get implied (r', w'); evaluate the government block for implied TR'; recompute the factor as the ratio of data mean income to model mean income.
4. **Update** by damped (convex-combination) iteration, $\Theta^{(i+1)} = \xi\, \Theta'^{(i)} + (1 - \xi)\, \Theta^{(i)}$ with damping $\xi \in (0, 1]$ (OG-Core default 0.4).
5. **Converge** when the maximum absolute distance between guessed and implied elements falls below tolerance, $\|\Theta'^{(i)} - \Theta^{(i)}\|_\infty <$ `mindist_SS` (default 10^{-9}), subject to an iteration cap (`maxiter`; the integration exposes this as `max_iter`, default 250 in the CLI).

Failures, when they occur, are almost always in step 2 on extreme reforms (tax functions implying near-confiscatory marginal rates somewhere on the grid) and manifest as non-convergence of the household root-finder; the documented mitigations are raising `max_iter` and falling back to pooled tax functions (Section 8).

3.7.2 Transition path iteration

Transition dynamics are solved by *time-path iteration* (TPI) in the tradition of Auerbach and Kotlikoff (1987), analysed for this model class by Evans and Phillips (2014). The object solved for is the entire equilibrium path over T periods (OG-Core defaults imply T on the order of $320 = 4S$ periods, long enough that the economy has effectively reached the new steady state well before the horizon ends). The algorithm:

1. **Terminal anchor.** Solve the post-reform steady state (Section 3.7.1); the path must land on it.
2. **Guess paths** $\{r_t, w_t, BQ_{j,t}, TR_t\}_{t=1}^{T}$, typically linear or exponential interpolations from initial-condition values to the terminal steady state.
3. **Household path solve.** Given the guessed price paths, solve every cohort's remaining-lifetime problem. This includes the $S - 1$ cohorts already alive at the reform date — a household aged s_0 at $t = 1$ solves a truncated problem of $2(E+S-s_0) + 1$ equations with its pre-reform asset holdings as initial condition — and the T cohorts born along the path, each solving a full $2S$-equation lifetime problem positioned diagonally in the (s, t) plane. These lifetime problems are independent given prices, and the implementation distributes them across Dask workers.

4. **Aggregate** along the time dimension to implied paths $\{\hat{K}_t, \hat{L}_t\}$, hence implied $\{r'_t, w'_t\}$ from the firm block, and implied $\{BQ'_{j,t}, TR'_t\}$ from adding-up and the closure rule.
5. **Update** the guessed paths by damped iteration with weight ξ_{TPI} (OG-Core default 0.4): $r_t^{(i+1)} = \xi_{TPI}\, r'_t + (1 - \xi_{TPI})\, r_t^{(i)}$, and likewise for the other paths.
6. **Converge** when the largest absolute gap between guessed and implied paths across all periods and path variables falls below `mindist_TPI` (default 10^{-5}, looser than the steady-state tolerance because the path system is far larger), subject to `maxiter`. A final feasibility check verifies goods-market clearing period by period (the resource-constraint residual, reported as `RC_error`).

One TPI solve costs on the order of $T\times$ a steady-state household solve per iteration and is orders of magnitude more expensive than the steady state alone — hours rather than tens of minutes on a laptop. In the PolicyEngine Macro CLI, only steady-state comparison is wired today; transition paths are available through the underlying `oguk` API (Section 5).

4 Tax functions from PolicyEngine microdata

The channel through which a statutory reform enters the general-equilibrium model is the tax-function estimation step of DeBacker et al. (2019). This is the PolicyEngine-specific part of the pipeline, and it is what distinguishes the psl-og member from a stylised-wedge OLG exercise: the households in Section 3.2 respond to a smooth summary of the *actual* UK tax and benefit code, computed reform-by-reform from microdata. This section describes the microdata pipeline end to end, writes out the functional forms in full — including all twelve parameters of the DEP specification and the estimation objective — and compares the forms on offer.

4.1 The microdata pipeline, from PolicyEngine UK to `txfunc`

The input data. The microdata substrate is PolicyEngine's *enhanced FRS*: the Department for Work and Pensions' Family Resources Survey (roughly 20,000 households per wave), enhanced in two ways. First, *imputation*: variables the FRS lacks or under-reports — capital income, wealth, consumption categories — are imputed from donor surveys (the HMRC Survey of Personal Incomes for the income distribution's upper tail, the Wealth and Assets Survey for asset holdings, the Living Costs and Food Survey for expenditure) using quantile-regression-forest models. Second, *reweighting*: household weights are re-optimised by gradient descent so that survey aggregates match a large set of administrative and official statistics (HMRC income-tax totals by band, DWP benefit caseloads, ONS demographics). The result is a household file that reproduces the published income distribution far better than the raw FRS — in particular in the top decile, which dominates income-tax revenue. The dataset is versioned and distributed through Hugging Face (gated; a `HUGGING_FACE_TOKEN` is required), and the integration caches it locally per year under the suite's cache directory.

Step 1: microsimulation. For each year y of the budget window and each policy P (baseline and reform), PolicyEngine UK (Woodruff et al., 2025) computes every household's full statutory position: income tax (allowances, bands, Scottish rates), National Insurance contributions, and the benefit system (Universal Credit with its taper, legacy benefits, child benefit and its high-income charge, pension credit). From the person-level output the pipeline extracts, for every adult i: labour income x_i (employment plus self-employment income), capital income y_i (savings interest, dividends, property and other capital income), total tax net of benefits T_i, and the survey weight w_i. The *effective tax rate* observation is $ETR_i = T_i/(x_i + y_i)$, computed on the sample with total income above a small floor (observations with near-zero income produce unbounded ratios and are dropped, as are extreme outliers by percentile trim).

Step 2: marginal rates by perturbation. Marginal tax rates are not read from statutory schedules — doing so would miss benefit withdrawal entirely. They are computed *numerically*: add a small increment Δ (£1) of employment income to every adult simultaneously, rerun the full microsimulation, and record

$$MTRx_i \;=\; 1 - \frac{\text{net income}_i(x_i + \Delta) - \text{net income}_i(x_i)}{\Delta}, \tag{28}$$

and analogously $MTRy_i$ perturbing capital income. Because the perturbation runs through the entire simulated system, the marginal-rate cloud automatically contains every kink and notch of the true schedule: the 55%-plus effective marginal rates where the Universal Credit taper overlaps income tax and NICs, the 60% band from personal-allowance withdrawal above £100,000, the child-benefit high-income charge, and the interactions among them. No hand-coded schedule could replicate this, and it is precisely this detail that the estimated functions summarise for the macro model.

Step 3: function fitting. For each year (and, if configured, each age cell), parametric functions $\tau^{etr}(x, y)$, $\tau^{mtrx}(x, y)$, $\tau^{mtry}(x, y)$ are fitted to the weighted scatter of micro observations by constrained weighted nonlinear least squares. Writing θ for the parameter vector of the chosen functional form and $\tau(\cdot; \theta)$ for the fitted function, the estimator for each rate type solves

$$\hat{\theta} \;=\; \arg\min_{\theta \in \Theta} \sum_i w_i \left[\tau_i - \tau(x_i, y_i; \theta)\right]^2, \tag{29}$$

where τ_i is the micro observation (ETR_i, $MTRx_i$ or $MTRy_i$), w_i the survey weight, and Θ imposes the shape constraints of the functional form (positivity of curvature parameters, $max \geq min$, $\phi \in [0, 1]$), so the fitted schedules are monotone and bounded by construction. OG-Core's `txfunc` module implements (29) with a global-then-local strategy (differential evolution or L-BFGS-B on transformed parameters, per `tax_func_type`), and reports the weighted sum of squared errors and observation count per fitted cell.

Step 4: embedding. The fitted parameters are handed to OG-Core, where the functions appear in the households' budget constraints and Euler equations (7)–(15). One subtlety makes the loop circular: the functions are estimated in *pounds*, but the model works in model units. The *factor* — the ratio of data average income to model average income — converts model incomes into pounds before the tax function is evaluated, and it is solved *jointly* with the steady state (it is one of the outer-loop unknowns of Section 3.7.1), because average model income itself depends on behaviour, which depends on taxes, which depend on the factor. Beyond the estimation sample's income support, rates are capped at the fitted asymptotic maxima, so extrapolation to the model's richest agents is conservative rather than explosive.

A reform score is then the comparison of two full general-equilibrium solves whose only primitive difference is the fitted tax-function parameters (plus any structural overrides).

4.2 The DEP functional form in full

OG-Core's richest specification is the *DEP* (DeBacker–Evans–Phillips) form: a bivariate function of labour income x and capital income y built as a Cobb–Douglas aggregate, with weight $\phi \in [0, 1]$, of two univariate ratio-of-polynomials schedules,

$$\tau(x, y) \;=\; \Big[\tau^x(x) + shift_x\Big]^{\phi} \Big[\tau^y(y) + shift_y\Big]^{1-\phi} + shift, \tag{30}$$

where each univariate component is a bounded, monotone ratio of quadratics,

$$\begin{aligned} \tau^x(x) \;&=\; \big(max_x - min_x\big)\,\frac{Ax^2 + Bx}{Ax^2 + Bx + 1} \;+\; min_x, \\ \tau^y(y) \;&=\; \big(max_y - min_y\big)\,\frac{Cy^2 + Dy}{Cy^2 + Dy + 1} \;+\; min_y, \end{aligned} \tag{31}$$

with $A, B, C, D > 0$. The twelve parameters per rate type are, explicitly:

#	Parameter	Constraint	Role
1	A	> 0	curvature of the labour-income schedule
2	B	> 0	slope of the labour-income schedule at low x
3	C	> 0	curvature of the capital-income schedule
4	D	> 0	slope of the capital-income schedule at low y
5	max_x	$> min_x$	asymptotic maximum rate in x
6	min_x	—	minimum rate in x (may be negative: net transfers)
7	max_y	$> min_y$	asymptotic maximum rate in y
8	min_y	—	minimum rate in y
9	$shift_x$	level shift	keeps the x-component positive before exponentiation
10	$shift_y$	level shift	keeps the y-component positive before exponentiation
11	$shift$	—	overall level shift (restores possibly negative rates)
12	ϕ	$\in [0, 1]$	Cobb–Douglas weight on labour vs. capital income

Each ratio $\frac{Ax^2+Bx}{Ax^2+Bx+1}$ rises smoothly from 0 at $x = 0$ to 1 as $x \to \infty$, so τ^x rises from min_x to max_x; the shifts guarantee the bracketed terms in (30) are strictly positive (so the Cobb–Douglas aggregate is well defined even when fitted minimum rates are negative, as they are for transfer-receiving low earners); and ϕ governs how much of the joint variation loads on labour versus capital income. The parameters are estimated separately for ETR, $MTRx$, and $MTRy$ — 36 parameters per year-age cell in total (DeBacker et al., 2019; DeBacker and Evans, 2024a). Figure 3 illustrates the resulting surface: rates rising smoothly in both income dimensions towards their asymptotic maxima.

4.3 Gouveia–Strauss as fitted, and the menu of forms

The framework also supports simpler forms on total income $z \equiv x + y$. The three-parameter Gouveia and Strauss (1994) schedule specifies tax *liability* $T(z) = \phi_0\big[z - (z^{-\phi_1} + \phi_2)^{-1/\phi_1}\big]$, so the effective rate is

$$\tau^{etr}(z) \;=\; \phi_0\Big[1 - \big(z^{-\phi_1} + \phi_2\big)^{-1/\phi_1}\, z^{-1}\Big], \tag{32}$$

and the marginal rate is the analytical derivative of the same liability function,

$$\tau^{mtr}(z) \;=\; T'(z) \;=\; \phi_0\Big[1 - \big(1 + \phi_2\, z^{\phi_1}\big)^{-1-\frac{1}{\phi_1}}\Big]. \tag{33}$$

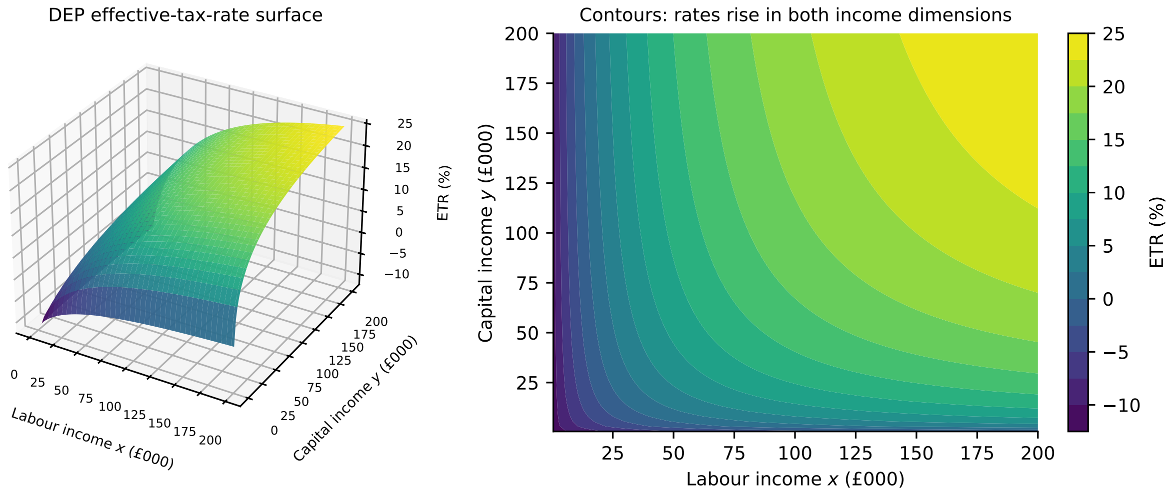

Figure 3: Illustrative DEP effective-tax-rate surface (30)–(31), drawn with parameter magnitudes in the range reported by DeBacker et al. (2019) (not a fitted UK surface). Left: the surface over labour income x and capital income y. Right: contours. The bounded ratio-of-quadratics components deliver rates that rise smoothly from a (possibly negative) minimum towards asymptotic maxima in each income dimension, so the model's richest agents face capped, not explosive, extrapolated rates.

Here $\phi_0 \in (0, 1)$ is the asymptotic top rate approached by both schedules, $\phi_1 > 0$ governs how fast progressivity sets in, and $\phi_2 > 0$ locates the schedule along the income axis. Because (33) is the exact derivative of the liability underlying (32), effective and marginal rates are *mutually consistent by construction* — an internal-consistency property the DEP form, whose three rate functions are estimated independently, does not share. Figure 4 plots a fitted-style pair. The third supported form is the log-linear progressivity function of Heathcote et al. (2017), $T(z) = z - \phi_0 z^{1-\phi_1}$ (so $\tau^{etr} = 1 - \phi_0 z^{-\phi_1}$), popular for its analytical tractability in optimal-tax work; and a piecewise-linear "mono" spline option exists for nonparametric fits on total income.

Comparing the forms. The trade-off across the menu is richness against robustness, summarised in Table 1. DEP is the only form that lets labour and capital income face different schedules — essential where the statutory system taxes them differently, as the UK does (separate dividend rates, savings allowances, NICs on labour income only) — but its 12-parameter nonconvex objective is the fragile step of the pipeline in thin cells, and its independently estimated ETR/MTR trio can be mutually inconsistent. GS and HSV collapse the joint distribution onto total income, cannot represent labour/capital asymmetry, but are fast, stable, and internally consistent. The OG-UK wiring deployed in PolicyEngine Macro fits the Gouveia–Strauss form to the PolicyEngine effective-rate microdata for each year of a three-year budget window, taking marginal-rate schedules as analytical derivatives; the richer DEP form remains available through OG-Core's `tax_func_type` option at the cost of a heavier and less robust estimation step.

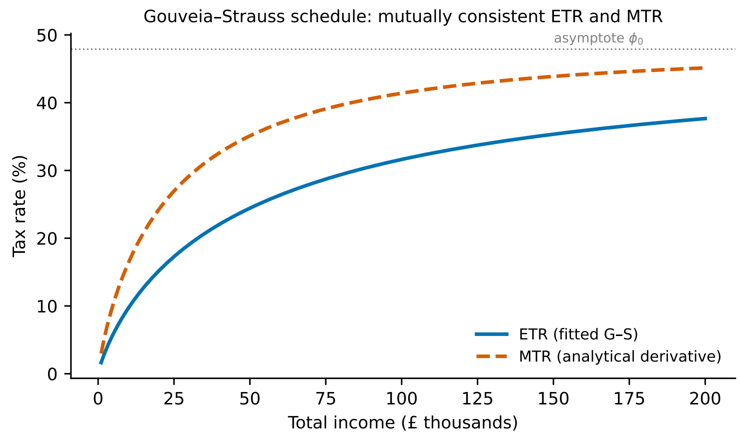

Figure 4: An illustrative Gouveia–Strauss fit (32)–(33) in the shape produced by the UK schedule: effective rates rising from near zero through the personal-allowance range towards the asymptotic maximum ϕ_0, with the marginal-rate schedule the analytical derivative of the same fitted liability function — so ETR and MTR cannot disagree, unlike independently estimated DEP surfaces.

Table 1: Tax functional forms supported by OG-Core's `txfunc`

Form	Parameters	Arguments	x/y asymmetry	ETR–MTR consistent	Deployed
DEP	12 per rate ($\times 3$)	(x, y)	yes	no (separate fits)	optional
Gouveia–Strauss	3	$x + y$	no	yes (analytic)	**default**
HSV log-linear	2	$x + y$	no	yes (analytic)	optional
Piecewise linear	knots	$x + y$	no	no	optional

4.4 Age pooling

Because ability profiles $e_{j,s}$ vary by age, tax functions can be estimated per age. The integration exposes three settings: `pooled` (one function for all ages — fastest and most stable, and the CLI default), `brackets` (one per age bracket, four groups split at the state pension age), and `each` (one per single year of age, $S = 80$ functions — closest to DeBacker et al. (2019)'s baseline but slowest and most fragile in thin cells). Pooling trades away age variation in effective schedules — pension-age interactions in particular — for robustness of the fit and of the equilibrium solve.

5 Implementation and integration

5.1 Software architecture

The stack has three layers, all open-source Python. At the bottom, **OG-Core** (`PSLmodels/OG-Core`) holds the country-agnostic theory of Section 3: the parameter apparatus, the `txfunc` estimation module, and the steady-state and TPI solvers, with transition-path lifetime problems distributed across Dask workers. Above it, **OG-UK** (`PSLmodels/OG-UK`, package `oguk`) supplies the UK calibration — demographics, macro parameters, and the PolicyEngine microsimulation hookup — and a deliberately small scoring API: `solve_steady_state(start_year, policy, ...)`, `run_transition_path(...)`, and `map_to_real_world(baseline, reform)`, which converts model-unit changes into current-price £bn using a GDP-anchored scale factor with levels tied to live ONS series (GDP, consumption, investment, government spending, the debt ratio) and HMRC total receipts, falling back to cached values offline. At the top, **policyengine-macro** (the PolicyEngine Macro integration package) wraps `oguk` in the suite's common adapters.

A reform is the suite's shared object — a flat {`parameter_path: value`} dictionary over PolicyEngine UK parameters, e.g. {`"gov.hmrc.income_tax.rates.uk[0].rate": 0.21`} for a 21% basic rate. The adapter builds a PolicyEngine `Policy` object from it, solves baseline and reform steady states, and returns levels, £bn changes, percent changes, and baseline/reform interest rates, together with the suite's common `ScoreResult` block (model class `olg-ge`, horizon `steady-state`) so that the same reform can be laid side by side with the OBR-emulator and microsimulation scores. Structural shocks outside statute — corporation tax (`cit_rate`), TFP (Z), demographics, spending shares — enter through `param_overrides` passed to the solvers, and can be combined with a statutory `Policy` in a single run.

5.2 The CLI: `pe-macro score --model og`

The command-line surface is two commands (plus the unified entry point):

```
pe-macro og-baseline
pe-macro og-score --reform
   '{"gov.hmrc.income_tax.rates.uk[0].rate": 0.21}'
pe-macro score --country uk --model og --reform '...'
```

Both use the fast configuration: pooled-age tax functions, a single representative sector, and steady-state comparison only. The baseline solve is cached at module level per `(start_year, max_iter)`, so repeated reform scores within a process pay for one baseline. The OG member is UK-only in the scorer today; passing another country is rejected with an explicit error.

5.3 Why local-only

The model is deliberately excluded from the hosted PolicyEngine Macro MCP server. A single steady-state solve — PolicyEngine microdata calibration plus the general-equilibrium fixed point — has been measured at over seventeen minutes on a laptop, a reform score needs two solves, and a transition path takes hours. That is incompatible with an interactive hosted tool whose other members answer in seconds, so the OG member runs only locally, and the suite's documentation states the runtime expectation up front rather than letting a hosted request time out.

5.4 Pinned-version fragility

The integration does not pin `oguk` itself — it is installed from the PSLmodels GitHub repository (version 0.3.2 at the time of writing) — but `oguk` in turn pins `policyengine-uk==2.88.0`. This is a real fragility, stated plainly: `policyengine-uk` $\geq$ 2.89 renamed the microdata dataset keys the calibration downloads (from `enhanced_frs_2023_24_<year>` to `populace_uk_*`), so mixing `oguk` with a newer `policyengine-uk` fails with a `KeyError` at calibration time. The adapter traps this failure mode and raises an actionable error naming both causes — the version mismatch, and the other common environment failure, a missing `HUGGING_FACE_TOKEN` for the gated enhanced-FRS dataset — rather than surfacing a bare stack trace. Until the pin is lifted upstream, the environment that runs OG-UK must hold `policyengine-uk==2.88.0`, which also

Table 2: Full parameter listing of the deployed OG-UK configuration

Parameter	Symbol	Value	Source / target
Model structure			
Active life span	S	80 (ages 20–99)	model structure
Childhood periods	E	20 (ages 0–19)	model structure
Lifetime-ability types	J	7	OG-Core default
Ability shares	λ_j	(.25,.25,.20,.10,.10,.09,.01)	OG-Core default (US tax-data estimation)
Ability profiles	$e_{j,s}$	80×7 matrix	OG-Core default; Figure 5
Production sectors	M	1 (option: 8)	ONS Blue Book supply-and-use tables when $M = 8$
Demographics			
Population by age	$\omega_{s,t}$	UK single-year profiles	UN World Population Prospects (United Nations, Department of Economic and Social Affairs, Population Division, 2024)
Fertility, mortality	f_s, ρ_s	UK profiles	UN WPP; Figure 1
Immigration rates	i_s	residual	reconciles successive UN population vectors
Preferences			
Coefficient of relative risk aversion	σ	1.5	OG-Core default
Discount factor	β	0.965 (fixed)	set consistent with the ONS household saving ratio; not estimated
Frisch elasticity	θ	0.35–0.4	Blundell et al. (2007); Peterman (2016)
Ellipse parameters	(b, υ)	(0.573, 2.856)	fitted to CFE at calibrated θ (Section 3.2.1)
Labour-disutility scale by age	χ^n_s	OG-Core profile	matched to US hours profile (inherited)
Bequest intensity	χ^b_j	OG-Core default	De Nardi (2004)-style warm glow
Time endowment	$\tilde{l}$	1.0	normalisation
Technology			
Capital share parameter	γ	0.35	retained OG-Core default (ONS factor shares would imply ≈ 0.40)
Elasticity of substitution	ε	1.0 (Cobb–Douglas)	deployed default
Depreciation rate	δ	0.065 / yr	ONS capital stocks and consumption of fixed capital
Long-run productivity growth	g_y	0.011 / yr	OBR potential-output growth (Office for Budget Responsibility, 2025a)
TFP level	Z	1.0	normalisation (shockable via overrides)
Government			
Corporate tax rate	τ^b	0.27 effective	UK corporation-tax schedule (structural override)
Tax depreciation rate	δ^τ	$= \delta$ baseline	capital-allowance reforms via overrides
Consumption tax rate	τ^c	VAT-anchored	HMRC effective VAT incidence
Transfers share	α_{tr}	GDP share	ONS/OBR expenditure data
Government consumption share	α_g	GDP share	ONS/OBR expenditure data
Long-run debt-to-GDP target	α_D	ONS-anchored	fiscal closure rule, Section 3.4
Closure instrument, phases	G; (T_{G1}, T_{G2}, ρ_G)	spending closure; (4, 20, 0.6)	UK-configured (OG-Core defaults: 20, 256, 0.1)
Government rate wedge	(τ_d, μ_d)	gilt-anchored	Bank of England yield data
World interest rate (open-economy features)	r^*	gilt-based	Bank of England yield data
Taxes on households			
Tax functions	$\tau^{etr}, \tau^{mtrx}, \tau^{mtry}$	estimated	PolicyEngine UK on enhanced FRS (Section 4)
Tax-function form	—	Gouveia–Strauss	deployed default; DEP optional
Tax-function age treatment	—	pooled	CLI default; `brackets`/`each` available

means the OG member scores reforms against a slightly older statute vintage than the suite's microsimulation member, which tracks the latest `policyengine` release. This divergence is a known, declared inconsistency of the current wiring, not a hidden one.

6 Calibration

OG-UK's calibration strategy follows the OG-Core template (DeBacker and Evans, 2024a; DeBacker et al., 2025): demographics from international and national statistical sources, macro parameters anchored to UK national accounts and official forecasts, preference parameters from the literature or matched to observed aggregates, and the tax detail supplied entirely by the estimated tax functions of Section 4. Table 2 collects the full parameter listing of the deployed configuration by block, with sources.

Three remarks on the table. First, the lifetime-ability apparatus (J, λ_j, $e_{j,s}$; Figure 5) is inherited from the OG-Core default calibration, which estimates ability profiles from tax microdata for the United States; a UK-specific re-estimation of $e_{j,s}$ from the enhanced FRS is an open calibration task, and results that depend on the upper tail of the ability distribution should be read with that in mind. Second, the discount factor is a fixed value ($\beta = 0.965$ annual, `estimate_beta=False`), chosen to be consistent with the ONS household saving ratio given the other parameters rather than estimated within the model. Third, the multi-sector option ($M = 8$ industries calibrated from ONS Blue Book supply-and-use tables by SIC section) exists for sector-level questions such as energy-price shocks, but the deployed scoring configuration is single-sector.

How these targets compare with the current official UK aggregates — and how the deployed parameter choices compare with the published OG-Core/OG-USA defaults and with the OBR's own UK OLG model — is consolidated in Section 7.

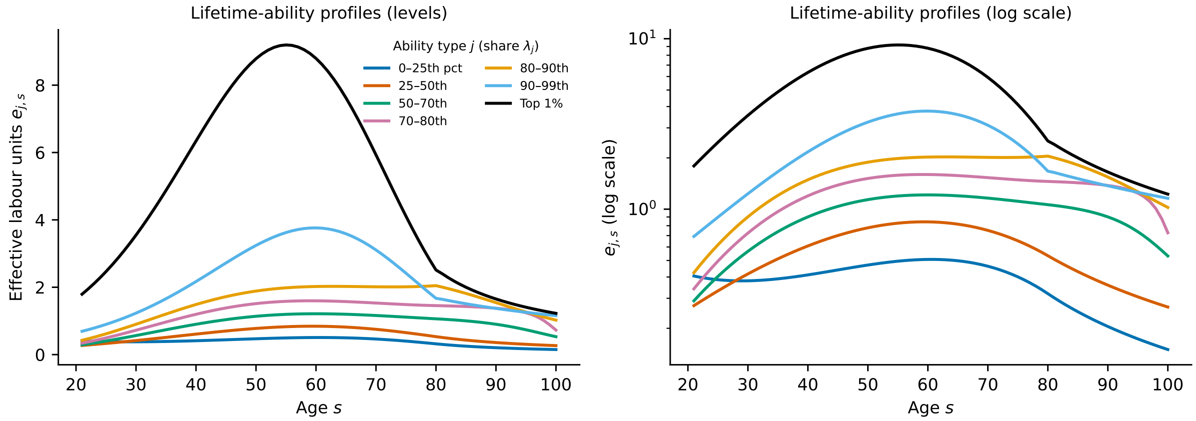

Figure 5: Lifetime-ability profiles $e_{j,s}$ for the $J = 7$ ability types, loaded from the OG-Core default `Specifications` object (levels, left; log scale, right). Type shares λ_j are (25, 25, 20, 10, 10, 9, 1)% from the bottom quartile to the top percentile. The top types' steep hump is what concentrates model income — and hence tax-function leverage — in a small share of agents. These profiles are estimated on *US* tax microdata and inherited by OG-UK; their UK re-estimation from the enhanced FRS is an open calibration task (Section 8).

7 Comparison with published results

Unlike the suite's OBR-emulator and SVAR members, the OLG member has no published ground truth to replicate (Section 8), so this section makes the honest comparisons that *are* available, in four parts: (i) the deployed calibration's targets against the current official UK aggregates; (ii) the deployed parameter choices against the OG-Core/OG-USA defaults and the published DeBacker et al. (2019) tax-function estimates the framework descends from, including the published fit statistics; (iii) the model's structure, calibration, and full published scenario results against the OBR's own UK overlapping-generations model, published as Working Paper 22 (Office for Budget Responsibility, 2025b); and (iv) the published record of OG-USA policy simulations, which brackets the plausible range of long-run responses the UK wiring should produce. Throughout, where a match is anchoring by construction rather than an independent validation, the tables say so.

7.1 Calibration targets against official UK aggregates

Because the £bn mapping is GDP-anchored (Section 5.1), the levels the model reports are anchored to ONS series by construction; the informative comparison is between the ratios the calibration *targets* and the corresponding official numbers. Table 3 reports the comparison against the most recently published official figures as of mid-2026, with the deviation stated explicitly, and Figure 6 plots the same comparison normalised by the official value.

The distinction in the treatment column matters for interpretation. Imposed quantities (depreciation, growth, the debt target) cannot disagree with the data; the genuinely over-identifying checks are the ones the steady state must *deliver* — the capital–output ratio and the interest rate consistent with the calibrated (γ, δ, β) triple. In the deployed fast configuration these emerge in an economically sensible range (a baseline steady-state real interest rate in low single digits, reported by `og-baseline` alongside the model-unit aggregates), but no systematic published

Table 3: Calibration targets and current official UK counterparts (as of mid-2026)

Quantity	Current official UK value	Source	Model treatment	Deviation
Labour share of income	$\approx$ 0.59–0.60, rising in 2024	ONS labour-share series / Blue Book 2025 (Office for National Statistics, 2025b)	$1-\gamma = 0.65$ (retained OG-Core $\gamma = 0.35$)	$\approx$ 5pp
Whole-economy net capital stock (produced-capital concept)	£5.6 trillion (2024), $\approx 2 \times$ GDP	ONS capital stocks bulletin (Office for National Statistics, 2025a)	K/Y emerges from $(\gamma, \delta, \beta, g_y)$; checked, not imposed[a]	concept-adjusted[a]
Depreciation / capital	$\approx$ 6–7% (CFC / net stock)	ONS capital stocks bulletin (Office for National Statistics, 2025a)	imposed, $\delta = 0.065$	within range
Government debt-to-GDP	95.1% (May 2026); 93.8% at end-FY 2025–26 (OBR March 2026 forecast: 94.4%)	ONS public sector finances (Office for National Statistics, 2026b)	imposed via closure target $\alpha_D = 0.95$	−0.1pp vs latest ONS
Potential growth	$\approx$ 1.1% / yr	OBR EFO (Office for Budget Responsibility, 2025a)	imposed, $g_y = 0.011$	$\approx$ 0
Household saving ratio	8.9% (2026 Q1, down from 9.6%)	ONS quarterly sector accounts (Office for National Statistics, 2026a)	$\beta = 0.965$ set consistent with it (fixed, not estimated)	consistent, not targeted

[a] The ONS whole-economy net stock includes dwellings ($\approx$ one third of the total), which the model's productive K does not correspond to one-for-one; the business-capital concept relevant to the model's K/Y is nearer 2, and the model's emergent ratio should be judged against that, not the headline 5.6/2.8.

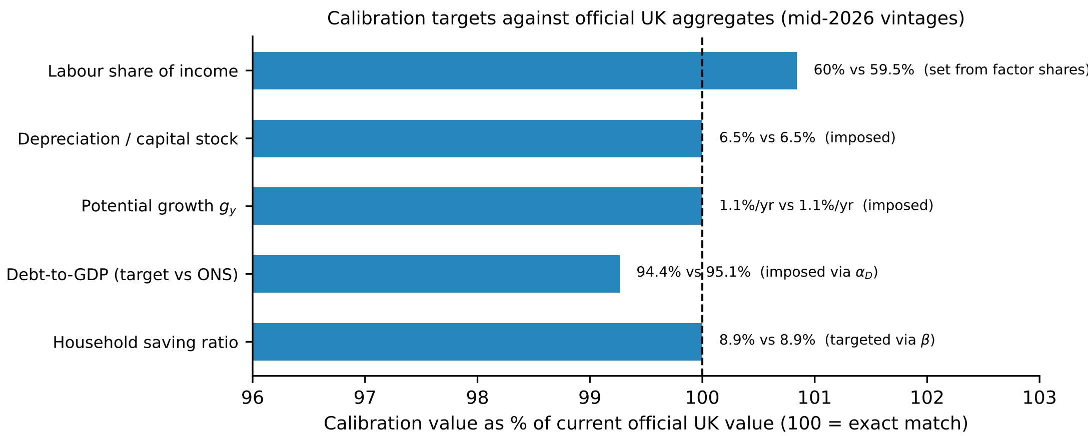

Figure 6: Calibration targets against current official UK aggregates, each expressed as a percentage of the official value (dashed line = exact match). Annotations report the underlying values and the model treatment: an exact match on an *imposed* or *targeted* quantity is anchoring, not validation — the over-identifying checks are the emergent ones (the capital–output ratio and the steady-state interest rate), for which no published reconciliation yet exists (Section 8).

reconciliation of achieved-versus-target moments exists for OG-UK 0.3.2, and this paper does not manufacture one: producing that reconciliation table from logged solves is the single most valuable next step for this member (Section 8). Two dating caveats on Table 3: the official column reports the figures current at writing (mid-2026), while the deployed calibration was frozen against slightly earlier vintages; and the saving ratio is volatile quarter to quarter, so the fixed β is judged against a smoothed level, not the latest print. The official column was re-verified against the releases available on 17 July 2026: the May 2026 public sector finances bulletin (debt at 95.1% of GDP, itself 0.7 percentage points above the OBR's March 2026 forecast of 94.4%) and the 2026 Q1 sector accounts (saving ratio 8.9%) remain the latest published prints, with the June 2026 finances bulletin due later in the month.

7.2 Parameter choices against the OG-Core/OG-USA published defaults

The second comparison is genealogical: OG-UK descends from OG-Core and its US calibration (DeBacker and Evans, 2024a,b; DeBacker et al., 2019), so every departure of the deployed UK configuration from the published defaults is a deliberate, explicable choice. Table 4 lays the two side by side.

The deviations follow a single logic: quantities that describe the *economy* (factor shares,

Table 4: Deployed OG-UK parameters against OG-Core/OG-USA published defaults

Parameter	OG-Core/OG-USA default	OG-UK deployed	Reason for deviation
Frisch elasticity θ	0.4	0.35–0.4	UK/US macro-Frisch evidence (Blundell et al., 2007; Peterman, 2016)
CRRA σ	1.5	1.5	retained: standard value, no UK-specific estimate imposed
Discount factor β (annual)	0.96	0.965	fixed value set consistent with the ONS household saving ratio (`estimate_beta=False`)
Depreciation δ (annual)	0.05	0.065	ONS consumption of fixed capital / net stock ratio
Capital share γ	0.35	0.35	retained OG-Core default; ONS factor shares would imply ≈ 0.40 (open re-anchoring task)
Productivity growth g_y (annual)	0.03	0.011	OBR potential-output growth: UK trend growth is far below the US default
Corporate tax rate τ^b	0.21 (US statutory)	0.27 effective	UK corporation-tax schedule
Consumption tax τ^c	0	VAT-anchored	the UK has a VAT; the US default abstracts from it
Ability types, shares J, λ_j	7; (.25, .25, .20, .10, .10, .09, .01)	unchanged	inherited; UK re-estimation is an open task (Section 8)
Ability profiles $e_{j,s}$	US tax-microdata estimates	unchanged	inherited; the flagged limitation of Figure 5
Tax-function form	DEP, age-specific (DeBacker et al., 2019)	Gouveia–Strauss, pooled	robustness over richness in the deployed fast configuration (Section 4.3)
Tax-function data	IRS/tax-calculator microdata	PolicyEngine UK on enhanced FRS	country substitution: the point of the OG-UK member

depreciation, trend growth, tax rates, the microdata substrate) are re-anchored to UK sources, while quantities that describe *preferences* (σ, the ability apparatus) are retained from the published US calibration unless UK evidence forces a change — with the retained US ability profiles the flagged weak point. The largest single numerical departure is trend growth: the US default $g_y = 0.03$ would be indefensible for the UK, where the OBR puts potential growth near 1.1% (Office for Budget Responsibility, 2025a), and since g_y scales the entire balanced-growth path this one substitution changes long-run levels more than any preference parameter.

On the tax-function side, the published reference point is the DEP fit that DeBacker et al. (2019) and the OG-USA documentation report for baseline US current law at age 42 in 2017: asymptotic maximum rates $max_x = 0.80$ ($max_x = 0.71$ for MTRx), fitted minimum rates of -0.14 to -0.17 (net transfer recipients), and a Cobb–Douglas income weight $\phi = 0.84$–0.96, estimated on 3,105 filer observations per rate. Figure 7 draws the schedules those published parameters imply and lays the paper's illustrative UK Gouveia–Strauss curve (the deployed functional form) alongside. The comparison is between *published US fits* and an *illustrative UK-shaped curve* — no fitted UK parameter vector is published for OG-UK 0.3.2, and this paper does not manufacture one — but it makes the structural point: both pipelines produce the same qualitative object, a smooth bounded schedule rising from near (or below) zero through the progressive range towards an asymptotic maximum, with the UK curve's faster early rise reflecting the personal-allowance taper and NICs arriving lower in the income distribution than the corresponding US thresholds.

Published fit statistics. DeBacker et al. (2019) also publish the fit diagnostics for their DEP estimation on US IRS-based microdata — the same weighted-NLLS objective (29) that the deployed pipeline minimises — and these give the only published yardstick for how well functions of this family fit microsimulation output. Table 5 reproduces the diagnostics for their illustrative age-43, year-2018 cell (a neighbouring cell to the age-42, 2017 parameter vector drawn in Figure 7; the published materials report different reference cells across vintages) under both 2017 law and the TCJA reform: roughly 3,500 weighted filer observations per rate type per cell, with root-mean-squared errors of 1.8–2.4 *percentage points* of tax rate, largest for the capital-income marginal rate (whose micro cloud is noisiest) and essentially unchanged between baseline and reform law. Two lessons carry over to the UK wiring. First, a well-behaved fit of this family leaves residual within-cell dispersion of around two percentage points — the parametric smoothing is a genuine compression, not an interpolation — so reforms whose effects live in fine distributional detail below that resolution should be scored with the microsimulation member, not this one. Second, no comparable fit table is yet published for the deployed UK Gouveia–Strauss estimation: OG-Core's `txfunc` reports the weighted SSE and observation count for every fitted cell at run time, but OG-UK 0.3.2 publishes no reference values, and producing that table

Table 5: Published DEP tax-function fit diagnostics, US age-43 cell, tax year 2018 (DeBacker et al., 2019)

	2017 law (baseline)			TCJA (reform)		
	ETR	MTRx	MTRy	ETR	MTRx	MTRy
Observations N	3,469	3,469	3,469	3,480	3,480	3,480
Weighted SSE	10,861	12,614	20,736	11,406	11,472	18,626
RMSE (pp of tax rate)	1.77	1.91	2.45	1.81	1.82	2.31

One of 3,600 estimated cells (3 rate types × 60 ages × 10 budget-window years × 2 policies); the authors present this cell as illustrative of the full set. Rates are measured in percentage points, so RMSE is directly interpretable as tax-rate dispersion around the fitted schedule.

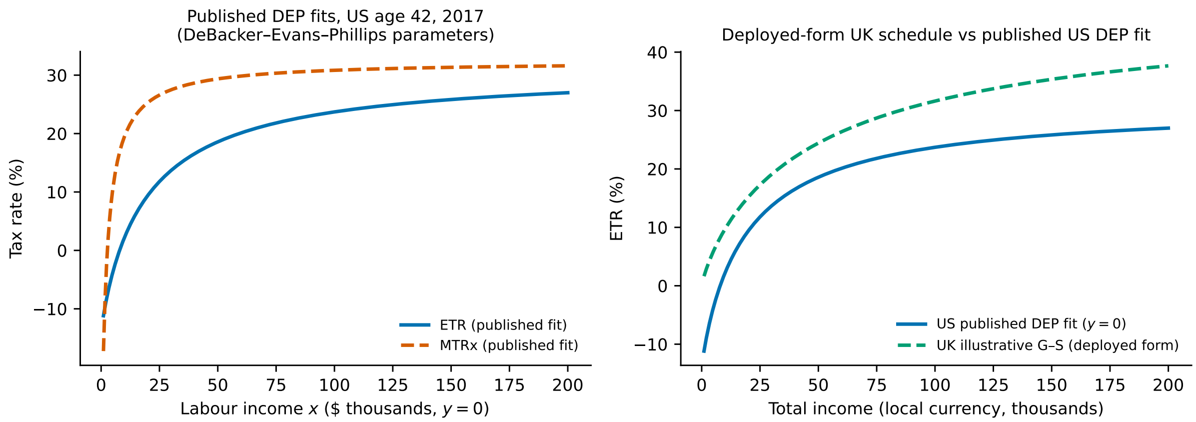

Figure 7: Left: effective and marginal labour-income tax-rate schedules implied by the DEP parameter vectors published for baseline US current law, age 42, 2017 (DeBacker et al., 2019; DeBacker and Evans, 2024b), evaluated at zero capital income. Right: the published US ETR curve against this paper's illustrative UK Gouveia–Strauss ETR (the deployed functional form; not a fitted UK vector, which OG-UK 0.3.2 does not publish). Axes are in local-currency thousands; the comparison is qualitative by design.

from logged UK estimations belongs to the same validation agenda as the achieved-versus-target moments (Section 8).

7.3 The OBR's UK OLG model as an external reference point

The closest published external reference is the OBR/HM Treasury UK OLG model of Office for Budget Responsibility (2025b): an independent codebase answering overlapping questions on the same economy. Table 6 compares the two models' structure and calibration using the parameter values published in that working paper's calibration tables.

The calibrations agree closely where both anchor to the same official sources (productivity growth 1.2% vs 1.1%; debt targets within a few points of current PSND) and differ where modelling philosophy differs: the OBR codes the statutory tax schedule directly and generates within-cohort inequality from earnings shocks, where OG-UK summarises the *simulated* tax system — benefit withdrawal included — as smooth estimated functions and carries inequality through deterministic ability types. Neither is dominant: the OBR's shock-based heterogeneity

Table 6: This paper's model (OG-UK as deployed) against the OBR's UK OLG (Office for Budget Responsibility, 2025b)

	OBR UK OLG (WP 22, April 2025)	OG-UK (this paper)
Cohorts	ages 20–100 ($J = 81$), annual	ages 20–99 ($S = 80$), annual
Within-cohort heterogeneity	idiosyncratic persistent + transitory earnings shocks	7 deterministic lifetime-ability types
Preferences	Cobb–Douglas CRRA, $\sigma_2 = 2$	CRRA in consumption, elliptical labour disutility, $\sigma = 1.5$
Discount factor	0.99 (calibrated to WAS asset profile)	fixed at 0.965, set consistent with ONS saving ratio
Frisch elasticity	$\approx$ 0.5 (0.3–0.5 to age 50, rising with age)	0.35–0.4 (constant target)
Capital share	0.33	0.35
Depreciation	5%	6.5%
Productivity growth	1.2% (March 2024 EFO)	1.1% (OBR potential growth)
Debt-to-GDP target	100%	ONS-anchored ($\approx$ 95%)
Tax detail	statutory rates and thresholds (March 2024) coded directly	smooth tax functions estimated from PolicyEngine microdata
Openness	closed baseline; open variant with $r^* = 5\%$	closed baseline; open-economy features gilt-anchored
Bequests	warm glow, target £107,000, received at 55–75	warm-glow intensity χ^b_j, accidental bequests redistributed
Consumption tax	effective VAT rate 9.3% (receipts / consumption)	VAT-anchored τ^c
Population growth	0.75%/yr (16+, March 2024 EFO)	UN WPP profiles, immigration residual
Solution method	value-function iteration on a discretised state space; core solve <2 minutes	Newton root-finding on Euler equations + time-path iteration; steady state $\approx$17 minutes

produces precautionary saving that the deterministic-types model cannot, while the microdata-estimated tax functions capture effective-rate detail (the UC taper, the £100,000 allowance withdrawal) that coded statutory schedules of the marginal-rate-and-threshold kind omit.

The OBR model's own fit to UK data. WP 22 also publishes the comparison between its calibrated baseline and the UK data (its Table 4.2), which is worth recording here because it shows how far even a carefully calibrated official OLG model sits from the aggregates: in the closed-economy baseline the consumption-to-GDP ratio is 74.6% against an outturn of 61.5% (the model has no net trade, housing, or government-consumption components of expenditure), private assets-to-GDP is 376% against 443% (the model under-generates peak retirement wealth), and the equilibrium real interest rate is 7.1% against an outturn near 3.9% — while income tax (9.9% vs 9.8% of GDP) and VAT (6.9% vs 7.3%) receipts match closely. The open-economy variant, which pins the interest rate at the assumed world rate of 5%, brings assets and debt interest nearer the data. This is the appropriate standard against which to judge the OG-UK member's emergent moments when its own reconciliation table is produced: close on the fiscal ratios that are effectively anchored, materially off on the equilibrium price and wealth moments that are not — a pattern to be expected of this model class, not a defect peculiar to either codebase.

Published scenario results. On results, WP 22 publishes one long-run scenario table (its Table 5.1), and its headline experiment is exactly the reform class this paper uses as its running example — a 1 percentage point change in the basic rate of income tax. Table 7 reproduces the published results in full for all four scenarios. For the basic-rate rise (A1), GDP per person falls 0.1%, total hours fall 0.2% (concentrated above age 50, peaking at about 10 minutes a week less for those aged 72), productivity *rises* 0.1% through composition as the least-productive workers reduce hours most, the consumption share falls 0.4 points, and income tax raises just over £8bn in real terms ($\approx$ 0.3% of GDP) — against the £6–8bn ($\approx$ 0.2% of GDP) short-run static yield in HMRC's ready reckoner, with WP 22 noting its own model overstates the yield somewhat because it under-represents incomes below the personal allowance. The cut (A2) is near-symmetric, with the residual budget item worsening 0.3% of GDP against the rise's 0.4% gain (a rounding artefact, per WP 22). The state-pension scenario (B) raises GDP per person 0.1% on a 0.2% labour-supply rise; the ageing scenario (C) lowers GDP per person 0.2% as a 0.5% labour-supply fall outweighs capital deepening, with assets-to-GDP up nearly 2 points and the real interest rate down 0.1 points.

Table 7: OBR UK OLG published long-run scenario results (WP 22, Table 5.1; differences from baseline)

	A1: basic rate +1pp	A2: basic rate −1pp	B: state-pension profile +1 year	C: ageing (survival +1pp)
GDP per capita (%)	−0.1	+0.1	+0.1	−0.2
Labour supply (%)	−0.2	+0.2	+0.2	−0.5
Labour productivity (%)	+0.1	−0.1	−0.1	+0.2
Real wage (%)	0.0	0.0	0.0	+0.3
Real interest rate (pp)	0.0	0.0	0.0	−0.1
Saving rate (pp)	0.0	0.0	0.0	−0.1
Consumption / GDP (pp)	−0.4	+0.4	−0.3	+0.1
Private assets / GDP (pp)	−0.2	−0.2	−0.2	+1.9
Residual budget item / GDP (pp)	+0.4	−0.3	+0.3	−0.1
Income tax / GDP (pp)	+0.3	−0.3	−0.1	0.0
Welfare spending / GDP (pp)	0.0	0.0	−0.3	+0.2

Closed-economy, full-anticipation, long-run (new steady state) results; WP 22 stresses these are "generic unit sensitivities, not policy costings". The A1 income-tax yield corresponds to just over £8bn in real terms.

Table 8: Published long-run GDP effects from OG-Core-family and comparable OLG simulations

Study	Model	Policy simulated	Long-run GDP effect	Budget-window effect
DeBacker et al. (2019)	OG-USA + Tax-Calculator	TCJA (2017 US tax act) vs prior law	+0.07% steady state[a]	+0.56% avg. 2018–27; +1.2% peak
Pomerleau et al. (2020)	OG-USA + Tax-Calculator	Biden 2020 platform (tax rises ≈ 1% of GDP)	−0.2%	−0.06% first decade; +0.07% second
Urban-Brookings Tax Policy Center (2020)	OG-USA + TPC microsim	Biden 2020 campaign proposals	positive after 2038 (crowding-in)[b]	−0.37%/yr avg. first decade
Office for Budget Responsibility (2025b)	OBR UK OLG (independent code)	UK basic rate ±1pp (≈ 0.3% of GDP)	∓0.1%	not published (long-run only)

[a] Preferred (age-varying DEP) specification; across the six tax-function specifications the steady-state effect spans +0.03% to +0.15%, with steady-state revenue about 3% lower — the choice of functional form moves the answer by roughly a factor of five even within one model. [b] Deficit-reducing packages raise long-run output in closed-economy OG models through reduced crowding-out, which is why the sign flips between horizons.

These are steady-state, closed-economy, full-anticipation numbers — the same experiment type as this member's CLI scoring — and they provide the natural external benchmark: an OG-UK basic-rate score whose signs match (output and hours down, revenue up by somewhat less than would be suggested by a static costing of the same statute) and whose magnitudes fall in the same sub-half-percent range would be corroborating; a materially larger response would point first at the Frisch calibration and the estimated MTR schedules. Beyond Table 7 and its baseline-fit table, WP 22 publishes no steady-state capital–output ratio or long-run elasticity schedule, so a full quantitative reconciliation is not yet possible from published numbers alone; producing the OG-UK side of this comparison from logged solves is part of the validation agenda of Section 8.

7.4 Published OG-USA policy simulations: the plausible range of long-run responses

The final comparison uses the published record of policy simulations run with OG-USA — the same OG-Core engine and tax-function pipeline as the UK wiring, applied by its own authors and by third-party institutions — to bracket the magnitude of long-run GDP responses this model family produces for tax reforms of stated size. Table 8 collects the published numbers alongside the OBR UK OLG basic-rate experiment.

Three regularities emerge (Figure 8), and they define the plausibility corridor for the UK wiring. First, *long-run GDP responses are small relative to the static size of the reform*: reforms

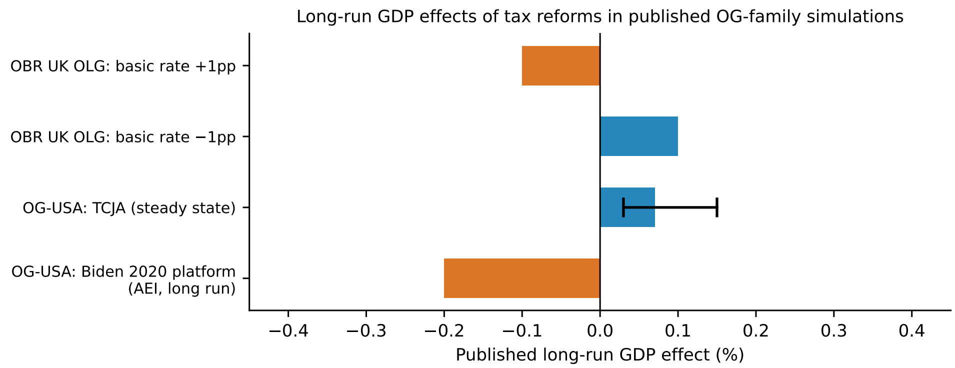

Figure 8: Published long-run GDP effects of tax reforms in OG-family simulations, as reported in the source documents (Table 8). The whisker on the TCJA bar is the published +0.03% to +0.15% spread across the six tax-function specifications of DeBacker et al. (2019) — the within-model sensitivity to the tax-function form alone. The TPC first-decade path (−0.37%/yr average, turning positive after 2038) is omitted because it is not a steady-state number.

moving of the order of 0.3–1% of GDP in revenue move long-run GDP by 0.1–0.4% across every published exercise, whether the model is OG-USA or the OBR's independent codebase. An OG-UK basic-rate score implying a multi-percent long-run GDP effect would therefore be presumptively wrong, whatever its internal convergence diagnostics say. Second, *the sign of the long-run effect can differ from the impact effect* for debt-changing reforms, because the closure rule and crowding-out dominate at long horizons — the TCJA's near-1% budget-window stimulus decays to +0.07% in the steady state, and the Biden package's first-decade drag turns positive once lower debt crowds capital back in. Steady-state-only scores (this member's CLI output) must therefore never be read as budget-window costings, exactly as Section 8 states. Third, *specification choices internal to the pipeline matter at the same order as the estimate itself*: the factor-of-five spread across tax-function forms in the TCJA exercise is the published quantification of a sensitivity the deployed UK configuration inherits, and it is the concrete reason this paper reports the deployed form (pooled Gouveia–Strauss) explicitly rather than treating the tax-function step as settled plumbing.

8 Limitations

This section states plainly what the psl-og member is not, in the spirit of the suite-wide rule that a model's status — replication, emulation, or structural counterfactual — is part of its result.

No ground truth. The OBR emulator and the SVAR members of the suite are validated against published figures; the OLG member has nothing comparable to replicate. Its outputs are structural counterfactuals: internally consistent general-equilibrium answers *conditional on* the model's preferences, technology, demographics, and closure rule. Two defensible OLG models can disagree materially on the long-run effect of the same reform — the OBR's own UK OLG model (Office for Budget Responsibility, 2025b) provides a natural cross-check, and side-by-side comparison against it is the closest available substitute for validation — so point estimates should be read as model-consistent magnitudes and signs, not forecasts.

Experimental-grade wiring. The integration is young. The version-pin fragility of Section 5.4 means the OG member currently runs against `policyengine-uk` 2.88.0 while the suite's microsimulation member tracks the latest release, so the two members can disagree about baseline statute; the gated enhanced-FRS microdata adds an access prerequisite; and the CI cannot exercise a seventeen-minute solve on every commit, so regressions in solve behaviour are caught by scheduled runs and by hand, not continuously.

Steady-state-only scoring today. The CLI compares long-run steady states. That is not a budget-window costing: it answers "where does the economy settle?", not "what happens in each of the next five fiscal years?". The transition-path solver exists and is reachable through the `oguk` API, but it is hours-long, Dask-distributed, and not yet wired into `pe-macro score`. Every scoring payload carries this caveat explicitly in its `ScoreResult` block.

Fast-configuration compromises. The deployed configuration pools tax functions across ages and uses one production sector. Pooling smooths over pension-age discontinuities in effective schedules; a single sector rules out compositional questions. Both dials can be turned up (`brackets/each`, `multi_sector=True`) at material runtime and stability cost — OLG solves can fail to converge on extreme reforms, and the documented mitigation (raise `max_iter`, fall back to pooled) is a workaround, not a guarantee.

Calibration vintage and inheritance. Calibration inputs carry vintages: demographic profiles from a UN release, depreciation and factor shares from a Blue Book vintage, potential growth from a particular EFO, and — most importantly — lifetime-ability profiles $e_{j,s}$ inherited from the US-estimated OG-Core defaults rather than re-estimated on UK microdata. No published achieved-versus-target moment reconciliation exists for the deployed version. Results that lean on the upper tail of the earnings distribution, or on fine age detail around retirement, are the least trustworthy outputs of the current calibration.

Model-class assumptions. Finally, the standard structural caveats apply: perfect foresight (no aggregate risk), rational optimising households, no involuntary unemployment, exogenous productivity growth, and a closure rule that mechanically stabilises debt. These are the assumptions that buy general-equilibrium discipline; conclusions that would reverse under plausible relaxations of them should not be attributed to the data.

9 Conclusion

The psl-og member gives the PolicyEngine Macro suite something none of its other members can: a general-equilibrium answer to long-run structural questions, in which a statutory reform changes wages, interest rates, saving, and labour supply through the optimising behaviour of eighty age cohorts. The model itself is community infrastructure — OG-Core (DeBacker and Evans, 2024a) and its OG-UK calibration (DeBacker et al., 2025) — and the suite's contribution is the wiring: the same flat PolicyEngine reform dictionary that drives the microsimulation and the OBR bridge also drives a full OLG solve, with reforms entering as tax functions estimated from PolicyEngine UK microdata in the manner of DeBacker et al. (2019), and results emerging in current-price £bn under a common, comparable score schema.

The member's constraints are structural and declared: it is local-only because a solve takes tens of minutes; it is steady-state-only in the CLI today; it rests on a pinned and slightly stale dependency pair; and it has no ground truth, so its numbers are counterfactuals, not forecasts. The near-term agenda follows directly: lift the `policyengine-uk` pin so all suite members share a statute vintage; publish an achieved-versus-target calibration reconciliation for the deployed version; re-estimate lifetime-ability profiles on UK microdata; wire the transition path into the scorer so the model can speak to fiscal-year profiles; and stand up a systematic comparison against the OBR's UK OLG model (Office for Budget Responsibility, 2025b) as the closest available external check. Each of these moves the member from experimental-grade wiring towards the standard the rest of the suite is held to; none changes what it fundamentally is — an open, inspectable structural model whose assumptions are on the table.

References

Auerbach, A. J. and Kotlikoff, L. J. (1987). *Dynamic Fiscal Policy*. Cambridge University Press, Cambridge.

Blundell, R., Costa Dias, M., Meghir, C., and Shaw, J. (2016). Female labor supply, human capital, and welfare reform. *Econometrica*, 84(5):1705–1753.

Blundell, R. and MaCurdy, T. (1999). Labor supply: A review of alternative approaches. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 3A, pages 1559–1695. Elsevier, Amsterdam.

Blundell, R., MaCurdy, T., and Meghir, C. (2007). Labor supply models: Unobserved heterogeneity, nonparticipation and dynamics. In Heckman, J. J. and Leamer, E. E., editors, *Handbook of Econometrics*, volume 6A, pages 4667–4775. Elsevier, Amsterdam.

Conesa, J. C., Kitao, S., and Krueger, D. (2009). Taxing capital? Not a bad idea after all! *American Economic Review*, 99(1):25–48.

De Nardi, M. (2004). Wealth inequality and intergenerational links. *Review of Economic Studies*, 71(3):743–768.

DeBacker, J. and Evans, R. W. (2024a). OG-Core: Documentation for the large-scale dynamic general equilibrium overlapping generations model. Policy Simulation Library, `https://pslmodels.github.io/OG-Core/`. Open-source model documentation, PSLmodels/OG-Core.

DeBacker, J. and Evans, R. W. (2024b). OG-USA: Documentation for the large-scale united states calibration of the OG-Core overlapping generations model. Policy Simulation Library, `https://pslmodels.github.io/OG-USA/`. Open-source calibration documentation, PSLmodels/OG-USA.

DeBacker, J., Evans, R. W., and Phillips, K. L. (2019). Integrating microsimulation models of tax policy into a DGE macroeconomic model. *Public Finance Review*, 47(2):207–275.

DeBacker, J., Evans, R. W., Woodruff, N., and Ghenis, M. (2025). OG-UK: An overlapping generations model to simulate fiscal policy in the united kingdom. Policy Simulation Library, https://github.com/PSLmodels/OG-UK. Open-source calibration of OG-Core; version 0.3.0.

Diamond, P. A. (1965). National debt in a neoclassical growth model. *American Economic Review*, 55(5):1126–1150.

Evans, R. W. and DeBacker, J. (2024). *Overlapping Generations Models for Policy Analysis: Theory and Computation*. Open Research Group, Inc. Online textbook, https://github.com/OpenRG/OGprimer.

Evans, R. W. and Phillips, K. L. (2014). OLG life cycle model transition paths: Alternate model forecast method. *Computational Economics*, 43(1):105–131.

Evans, R. W. and Phillips, K. L. (2018). Advantages of an ellipse when modeling leisure utility. *Computational Economics*, 51(3):513–533.

Gouveia, M. and Strauss, R. P. (1994). Effective federal individual income tax functions: An exploratory empirical analysis. *National Tax Journal*, 47(2):317–339.

Heathcote, J., Storesletten, K., and Violante, G. L. (2017). Optimal tax progressivity: An analytical framework. *Quarterly Journal of Economics*, 132(4):1693–1754.

İmrohoroğlu, A., İmrohoroğlu, S., and Joines, D. H. (1995). A life cycle analysis of social security. *Economic Theory*, 6(1):83–114.

Kotlikoff, L. J., Smetters, K., and Walliser, J. (2007). Mitigating america's demographic dilemma by pre-funding social security. *Journal of Monetary Economics*, 54(2):247–266.

Nishiyama, S. and Smetters, K. (2007). Does social security privatization produce efficiency gains? *Quarterly Journal of Economics*, 122(4):1677–1719.

Office for Budget Responsibility (2025a). Economic and fiscal outlook. Technical report, Office for Budget Responsibility, London.

Office for Budget Responsibility (2025b). The UK overlapping generations model (UK OLG). Working Paper 22, Office for Budget Responsibility, London.

Office for National Statistics (2024). UK national accounts, the blue book. https://www.ons.gov.uk/.

Office for National Statistics (2025a). Capital stocks and fixed capital consumption, UK: 2025. https://www.ons.gov.uk/economy/nationalaccounts/uksectoraccounts/bulletins/capitalstocksconsumptionoffixedcapital/2025.

Office for National Statistics (2025b). Labour share of income: Whole economy, SA: Percentage: UK (series FZLN); and UK national accounts, the blue book 2025. https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/labourproductivity/timeseries/fzln/ucst.

Office for National Statistics (2026a). Households' saving ratio (series DGD8), quarterly sector accounts, UK: Quarter 1 2026. https://www.ons.gov.uk/economy/grossdomesticproductgdp/timeseries/dgd8/ukea.

Office for National Statistics (2026b). Public sector finances, UK: May 2026. https://www.ons.gov.uk/economy/governmentpublicsectorandtaxes/publicsectorfinance/bulletins/publicsectorfinances/may2026.

Peterman, W. B. (2016). Reconciling micro and macro estimates of the Frisch labor supply elasticity. *Economic Inquiry*, 54(1):100–120.

Policy Simulation Library (2024a). The Policy Simulation Library: A catalog of open-source policy models. https://pslmodels.org/.

Policy Simulation Library (2024b). Tax-Calculator: An open-source microsimulation model of the US federal individual income and payroll tax system. https://github.com/PSLmodels/Tax-Calculator.

Pomerleau, K., DeBacker, J., and Evans, R. W. (2020). An analysis of Joe Biden's tax proposals. Technical report, American Enterprise Institute, Washington, DC. https://www.aei.org/research-products/report/an-analysis-of-joe-bidens-tax-proposals/.

Samuelson, P. A. (1958). An exact consumption-loan model of interest with or without the social contrivance of money. *Journal of Political Economy*, 66(6):467–482.

United Nations, Department of Economic and Social Affairs, Population Division (2024). World population prospects 2024. https://population.un.org/wpp/.

Urban-Brookings Tax Policy Center (2020). Macroeconomic analysis of former Vice President Biden's tax proposals, updated with OG-USA results. Technical report, Urban Institute & Brookings Institution, Washington, DC. https://taxpolicycenter.org/publications/macroeconomic-analysis-former-vice-president-bidens-tax-proposals.

Woodruff, N., Ghenis, M., and PolicyEngine contributors (2025). PolicyEngine UK: An open-source microsimulation model of the UK tax and benefit system. https://github.com/PolicyEngine/policyengine-uk. Version 2.88.0.

Yaari, M. E. (1965). Uncertain lifetime, life insurance, and the theory of the consumer. *Review of Economic Studies*, 32(2):137–150.

Zodrow, G. R. and Diamond, J. W. (2013). Dynamic overlapping generations computable general equilibrium models and the analysis of tax policy. In Dixon, P. B. and Jorgenson, D. W., editors, *Handbook of Computable General Equilibrium Modeling*, volume 1, pages 743–813. Elsevier, Amsterdam.

A Parameter glossary

Table 9 collects every symbol used in the paper, its meaning, and where it is set.

Table 9: Glossary of model symbols

Symbol	Meaning	Type	Where set
E, S	childhood periods; active life span	structure	fixed (20, 80)
J, λ_j	number of ability types; type shares	structure	OG-Core default (7; shares in Table 2)
$e_{j,s}$	effective labour units by type and age	exogenous profile	OG-Core default matrix (Figure 5)
$\omega_{s,t}$, $\hat{\omega}_{s,t}$	population by age; normalised shares	exogenous path	UN WPP + laws of motion (1)–(2)
f_s, ρ_s, i_s	fertility, mortality hazard, immigration rate	exogenous profiles	UN WPP; i_s residual
$g_{n,t}$, $\bar{g}_n$	population growth rate; stationary value	derived	from demographic block
g_y	labour-augmenting productivity growth	calibrated	OBR potential growth
$c_{j,s,t}$, $n_{j,s,t}$, $b_{j,s,t}$	consumption, labour, assets	endogenous	household problem
σ	coefficient of relative risk aversion	calibrated	1.5
β	discount factor	matched	ONS saving ratio
χ^n_s, χ^b_j	labour-disutility scale; bequest intensity	calibrated	OG-Core profiles
$v(n)$; b, υ, $\tilde{l}$	elliptical disutility; its parameters	fitted	CFE-matching (Section 3.2.1)
θ	Frisch elasticity of labour supply	calibrated	0.35–0.4
x, y	labour income; capital income	endogenous	$x = we_{j,s}n$, $y = r_pb$
τ^{etr}, τ^{mtrx}, τ^{mtry}	effective and marginal tax-rate functions	estimated	PolicyEngine microdata (Section 4)
ϕ_0, ϕ_1, ϕ_2	Gouveia–Strauss parameters	estimated	`txfunc` NLLS (29)
A–D, max/min, $shifts$, ϕ	DEP parameters (12 per rate)	estimated	`txfunc` NLLS (Section 4.2)
factor	model-to-data income scaling	endogenous	solved with steady state
bq, BQ_j	received bequests; aggregate by type	endogenous	accidental-bequest redistribution
tr, TR_t, α_{tr}	transfers; aggregate; GDP share	rule	government block
G_t, α_g	public consumption; GDP share	rule / closure	government block
D_t, α_D	government debt; long-run debt/GDP target	rule / closure	fiscal closure (24)
T_{G1}, T_{G2}, ρ_G	closure phase boundaries; smoothing weight	configured: 4, 20, 0.6	UK-configured (OG-Core defaults 20, 256, 0.1)
r_t, w_t	return on capital; wage	endogenous	firm FOCs (19)–(20)
$r_{gov,t}$; τ_d, μ_d	government rate; its wedge parameters	calibrated	gilt-anchored (22)
$r_{p,t}$	household portfolio return	derived	(8)
K_t, L_t, Y_t, C_t, I_t	aggregates	endogenous	market clearing
Z_t, γ, ε, δ, δ^τ	TFP; capital share; substitution elasticity; economic and tax depreciation	calibrated	technology and tax blocks
τ^b_t, τ^c_t	corporate and consumption tax rates	calibrated / override	government block
ξ, ξ_{TPI}	damping weights (SS, TPI)	solver	OG-Core defaults (0.4)
`mindist_SS`, `mindist_TPI`	convergence tolerances	solver	10^{-9}, 10^{-5}

B The OG-Core API surface used by the integration

The integration touches a deliberately small slice of the two underlying packages. From **OG-Core** (`ogcore`, version 0.x series):

- `ogcore.parameters.Specifications` — the parameter container. The integration constructs one per run, applies OG-UK's default parameter JSON (`oguk_default_parameters.json`), and layers any `param_overrides` (e.g. `cit_rate`, TFP, spending shares) via `update_specifications`.
- `ogcore.txfunc` — tax-function estimation: the NLLS objective (29) over the configured `tax_func_type` (`GS` deployed; `DEP`, `HSV`, `mono` available).
- `ogcore.SS.run_SS` — the steady-state fixed point of Section 3.7.1; returns the steady-state dictionary (aggregates, prices, policies by (j, s), the factor).
- `ogcore.TPI.run_TPI` — time-path iteration of Section 3.7.2; reachable through `oguk` but not wired into the CLI scorer.
- `ogcore.execute.runner` — the orchestration entry point OG-UK's API wraps (parameter setup, optional tax-function estimation, SS and/or TPI, output pickling to `OUTPUT/SS/SS_vars.pkl` and `OUTPUT/TPI/TPI_vars.pkl`).
- Supporting modules used indirectly: `demographics` (UN data retrieval and the population

machinery of Section 3.1), `elliptical_u_est` (the (b, υ) fit), **`household`**, **`firm`**, **`fiscal`**, **`tax`**, **`aggregates`** (the equation blocks), and `output_plots`/`output_tables` (diagnostics).

From **OG-UK** (`oguk` 0.3.2, installed from GitHub):

- `oguk.get_micro_data` — runs PolicyEngine UK over the enhanced FRS (cached per year under the suite cache directory) and produces the $(x_i, y_i, ETR_i, MTRx_i, MTRy_i, w_i)$ observations of Section 4.1, including the £1-perturbation marginal rates.

- `oguk.calibrate.Calibration` — assembles demographics, macro parameters, and estimated tax functions into a `Specifications`-compatible parameter set.

- `oguk.api.solve_steady_state` and `oguk.api.run_transition_path` — the scoring API, taking `start_year`, a `policy`, and solver options; `policy` is a PolicyEngine `Policy` object built from the suite's flat reform dictionary.

- `oguk.api.map_to_real_world(baseline, reform)` — converts model-unit changes into current-price £bn using a GDP-anchored scale factor with levels tied to live ONS series and HMRC receipts, falling back to cached values offline.

The **policyengine-macro** adapter contributes the suite contract on top: reform-dictionary validation, the module-level baseline cache keyed on `(start_year, max_iter)`, the `ScoreResult` block (model class `olg-ge`, horizon `steady-state`), and the actionable error mapping for the two known environment failures (version mismatch; missing Hugging Face token).

C Reproduction recipe

The following recipe reproduces the deployed scoring environment exactly, on a machine with Python 3.10+ and ≥16 GB of memory. Versions are pinned deliberately (Section 5.4); do not substitute newer ones.

1. **Environment.** Create a fresh virtual environment and install the pinned pair plus the integration:

   ```
   python -m venv .venv && source .venv/bin/activate
   pip install policyengine-uk==2.88.0
   pip install git+https://github.com/PSLmodels/OG-UK # oguk 0.3.2
   pip install policyengine-macro
   ```

   The `oguk` dependency resolution pulls a compatible `ogcore` automatically; `policyengine-uk` must be 2.88.0 exactly — versions $\geq$ 2.89 renamed the microdata dataset keys (`enhanced_frs_2023_24_*` → `populace_uk_*`) and fail at calibration time with a `KeyError`.

2. **Data access.** Export a Hugging Face token with access to the gated PolicyEngine enhanced-FRS dataset:

   ```
   export HUGGING_FACE_TOKEN=hf_...
   ```

   On first run the calibration downloads one `.h5` file per budget-window year into the suite cache directory (`~/.cache/policyengine-macro/policyengine-data/`); subsequent runs are offline.

3. **Baseline.** `pe-macro og-baseline` solves and reports the baseline steady state (model-unit aggregates, the interest rate, and the £bn anchoring). Expect ≈17+ minutes on a laptop: several minutes of microsimulation and tax-function estimation, the remainder in the general-equilibrium fixed point.

4. **A reform score.**

   ```
   pe-macro score --country uk --model og --reform \
      '{"gov.hmrc.income_tax.rates.uk[0].rate": 0.21}'
   ```

   Two solves are required (baseline and reform); within one process the baseline is cached, so a batch of reforms pays for the baseline once.

5. **Determinism and outputs.** Solves are deterministic given versions and data vintage: the pipeline involves no simulation randomness, so re-runs reproduce results to solver tolerance (10^{-9} in the steady-state outer loop). Raw solver output is pickled under `OUTPUT/` in the working directory (`SS_vars.pkl`; `TPI_vars.pkl` when a transition path is run via the `oguk` API), and the CLI's JSON `ScoreResult` carries the levels, £bn changes, percent changes, and the status caveats of Section 8.

6. **Figures.** The illustrative figures in this paper regenerate with `python figures/make_figures.py` from the paper directory; they load only the UN demographic CSVs shipped with `oguk`, the OG-Core default `Specifications`, and published parameter values — no equilibrium solve is required.