Supplemental Appendix to

"Calibration as a Channel for Discretization Error"

Orhan Torul

Boğaziçi University

August 27, 2026

This supplement collects the appendices to the main text: the computational implementation (Appendix A), empirical targets and sources (Appendix B), additional properties of the case-study economy (Appendix C), the economic sensitivities and extensions (Appendix D), the majority-rule tournament (Appendix E), the levels-versus-logs comparison (Appendix F), and phased reform horizons (Appendix G). Numbered references to sections, tables and equations outside the A–G range point to the main text.

Appendices

A Computational Implementation

Section 4.3 sets out the two structural features that make the household maximization separable and the grid design that follows from where the stationary distribution puts its mass. This appendix records the remaining implementation detail.

The outer loop iterates on the aggregate state $(\bar{H}, \bar{E}, \bar{T}r)$ with damping 0.3 to a tolerance of 3×10^{-5}; the inner value-function iteration uses the MacQueen and Porteus (1975) error bounds and is converged several orders of magnitude tighter at every outer step. The production discretization is $N_h = 240$ power-spaced nodes per equation (16), $N_e = 60$ education nodes on $[0, 2]$ with $e = 0$ flagged as the public-school option, $N_l = 40$ labor nodes on $[0.05, 0.50]$, $N_\xi = 19$ Tauchen (1986) nodes on $\log \xi$ spanning $\pm 3\sigma_\xi$, and $N_\varepsilon = 11$ Tauchen nodes on the log earnings shock spanning $\pm 3\sigma_\varepsilon$, rescaled so the discretized variance is exact and recentred so $\mathbb{E}[\exp \varepsilon] = 1$. Values of h' are interpolated linearly onto the h-grid. Kopecky and Suen (2010) show that the Tauchen discretization degrades for highly persistent processes, where the Rouwenhorst method does not. Neither process here is in that range: the earnings shock is i.i.d. and the benchmark ability process sets $\kappa = 0$, with $\kappa = 0.4$ in the robustness exercise of Appendix D.4. Table 9 varies N_ξ and N_ε alongside the other dimensions.

Because equation (5) makes income strictly positive, the shock grid need not be truncated to keep consumption positive, and the numerical guard on c does not bind at either calibrated regime, at any point of the policy grids used in Sections 5 and D, or in the non-labor income extension. The solver reports the probability mass at the guard and it is zero throughout.

At the coarse discretization the outer loop does not reach a point fixed point near the turning point of the ∪-shape, where the policy alternates between adjacent education nodes as the aggregate state moves, producing a limit cycle of order 10^{-3} in $\bar{H}$. This is a symptom of the same under-resolution that Section 5 documents, and it does not occur at the production grid, where the residual falls below the outer tolerance.

Transitions are solved by backward induction on V along the announced policy path with terminal condition $V_T = V(\text{terminal regime})$, coupled to forward iteration on μ_t from the status-quo distribution, with damped fixed-point iteration on the aggregate path. The benchmark horizon is $T = 25$ generations. The convergence study in Table 9 re-solves the transition and recomputes consumption-equivalent welfare at every discretization, so the welfare number reported there is subject to the same check as the stationary moments.

The Aiyagari economy of Section 7.1: The second application shares no code with the model above. The household problem is solved by the endogenous grid method of Carroll (2006); the invariant distribution is computed by iterating a sparse lottery operator on the asset grid to a tolerance of 10^{-11}; the capital market is cleared by bisection on aggregate capital; and the discount factor is calibrated by bisection on the capital–output ratio.

The converged configuration is 600 power-spaced asset nodes on $[0, 100]$, in units of mean labor income. Refining to 1,200 nodes, raising the spacing power from three to four, doubling the upper bound, or going to 2,000 nodes each moves the reported conclusion by at most 0.001 percentage points and the capital–output ratio by at most 0.0002, so the fine column of Table 13 is converged for the purpose it serves.

The replication archive contains the full pipeline. Every numeric table in the paper is emitted directly from computed output by a single script, which refuses to build a table from results produced at superseded parameter values; no number in the paper is transcribed by hand.

B Empirical Targets and Sources

This appendix documents the construction and sources of the empirical moments reported in Sections 4.1 and 4.4.4.

B.1 Cross-country macro variables

Tax-to-GDP ratio: OECD *Revenue Statistics*, 2018 release. Total tax revenue as a share of GDP, including social-security contributions per the OECD definition. The Europe-13 average is unweighted across the thirteen countries. Ireland's tax-to-GDP ratio understates its tax revenue per resident because of multinational tax-base distortion of the GDP denominator; an alternative computation using the modified-GNI denominator raises the Irish figure by approximately fifteen percentage points. None of the cross-country patterns reported in this paper depends materially on the treatment of Ireland.

Public share of education and health expenditure: OECD *Education at a Glance* (Table C3.2, 2020 release) for education; OECD *Health at a Glance* (Figure 7.11, 2017 release) for health. Switzerland's health-care figure of 30.5% in the 2017 OECD release reflects an exclusion of compulsory private health insurance from the "public" share that is not applied symmetrically to other countries; under the OECD's *Health at a Glance 2025* (SHA-2011-aligned) definition, the corresponding figure is 68.0%. The latter is the value reported in Section 4.1, for cross-country comparability. Switzerland is omitted from the education panel because the source does not report a value; the Europe-13 average for education is therefore over the remaining twelve countries.

Annual hours per worker: OECD productivity database, 2015 column. Year-on-year variation in this series is small for the Europe-13 sample (typically below 2% between 2015 and 2019), so the 2015 figure is taken as the pre-COVID reference.

Gini coefficients: OECD Income Distribution Database, December 2025 release. Disposable-income Gini values are uniformly 2019; market-income Gini values are taken from the closest pre-COVID year available for each country (2018–2022 depending on country); the country-by-country vintages are listed in the replication archive.

GDP per capita at PPP: World Bank World Development Indicators, indicator NY.GDP.PCAP.PP.CD, 2018 value. Same caveat as for tax-to-GDP regarding the Irish denominator.

B.2 World Values Survey item E040

The luck-leaning belief share $\bar{p}_c$ for country c is constructed as

$$\bar{p}_c = \frac{\sum_{w\in W_c} N_{cw}\, p_{cw}}{\sum_{w\in W_c} N_{cw}}, \qquad p_{cw} = \frac{\sum_i s_i\, \mathbf{1}\{r_i \geq 6\}}{\sum_i s_i},$$

where W_c is the set of waves available for country c, N_{cw} is the unweighted sample size, r_i is the respondent's E040 score on the 1-to-10 scale, and s_i is the within-wave post-stratification weight (variable S017 in the integrated WVS-EVS file). The Europe-13 average is unweighted across countries.

The set W_c pools waves WVS5 (2005–2009) and WVS6 (2010–2013) where both are available. Wave WVS7 (2017–2022) is excluded because it spans the COVID structural break. Three countries (Belgium, Denmark, Ireland) lack pre-COVID coverage of E040 and fall back to the only available reading, the European Values Study wave of 1990. Denmark's reading of 56.3% is the largest in the sample and is best treated as a high-noise observation given the vintage; the Europe-13 average is robust to dropping Denmark and falls modestly to 34.7%.

B.3 Residual-earnings estimation on the PSID

As a direct check on the U.S. anchor I estimate the standard residual-earnings decomposition on the Panel Study of Income Dynamics, using the harmonized PSID-SHELF release (Daumler et al., 2025).[1] The permanent and transitory variance components are $\hat{\sigma}^2_{\text{perm}} = 0.047$ (s.e. 0.001) and $\hat{\sigma}^2_{\text{trans}} = 0.039$ (s.e. 0.001), against the Krueger et al. (2010) Table 7.A U.S. figures of 0.028 and 0.060, roughly two-thirds higher and a third lower respectively, and not a close match. The pattern is what one would expect if earnings absorb hours-driven dispersion that a wage-based measure nets out, though I do not verify that decomposition here.

That same panel also delivers a direct empirical counterpart to the object in equation (14), which Appendix C.3 uses to discipline the model's luck share from outside. The unexplained share of the cross-sectional variance of log earnings, residual variance over total variance, is 0.705 (s.e. 0.006) under the controls above, and 0.703 (s.e. 0.006) when the control set is enriched with age-education interactions and five-year cohort effects. This is an *upper bound* on the model's ι_l rather than an estimate of it, since what an econometrician cannot explain is not what a household cannot foresee, and a household acts on career prospects and ability that the regressors do not contain. Netting out the person-fixed component of the residual, which the household knows by construction,

[1] Male and female reference persons aged 25–60 with positive real labor earnings; annual waves 1968–1997, 126,240 person-years across 15,267 individuals. The first stage regresses log real labor earnings on a fourth-order age polynomial, years of schooling and its square, sex, race, region, and year effects. Standard errors are from a person-level cluster bootstrap with $B = 200$.

gives a tighter upper bound of 0.287 (s.e. 0.004). Both bounds are stable across the sample: the unexplained share is 0.61, 0.72 and 0.73 over 1968–77, 1978–87 and 1988–97.

C Additional Properties of the Case-Study Economy

This appendix collects properties of the calibrated economy that support the main text but are not needed for the methodological argument: the household decision rules, the distributional detail behind the scorecard, the nested luck-share decompositions and their empirical bounds, and the Atkinson aggregation of the welfare results.

C.1 Household decision rules

The behavior behind the ∪-shape is visible in the education policy function and is unchanged from Torul (2020), so I describe it rather than re-plotting it. The public-school region, the set of states at which $e^* = 0$, widens monotonically in the size of government, from nearly empty at laissez-faire to almost the whole state space at $\bar{T}/\bar{Y} = 0.20$. Conditional on choosing private schooling, households with higher inborn ability spend less for given h, because the multiplicative law of motion raises their continuation value at any h' and the income effect dominates under log utility. What earnings risk adds is a second reason for the public option to be chosen: because the shock is realized after schooling is committed, a household near the margin values the fact that public provision does not have to be financed out of a realization it cannot foresee.

C.2 Distributional properties

Table C.1 reports Gini coefficients along the two policy dimensions. Three patterns matter for what follows.

Table C.1: Gini coefficients across the two fiscal instruments

Panel A: size of government, at $\tau_p = \tau_p^{US}$							
$\bar{T}/\bar{Y}$	$G(h)$	$G(e)$	$G(e_{\text{tot}})$	$G(l)$	$G(y)$	$G(\tilde{y})$	$G(c)$
0.0000	0.482	0.272	0.272	0.002	0.315	0.274	0.274
0.0250	0.483	0.293	0.272	0.012	0.319	0.275	0.272
0.0500	0.492	0.414	0.313	0.021	0.328	0.280	0.267
0.0750	0.496	0.646	0.330	0.030	0.334	0.282	0.259
0.0974	0.490	0.808	0.248	0.027	0.329	0.275	0.255
0.1166	0.482	0.896	0.155	0.022	0.322	0.267	0.253
0.1500	0.473	0.969	0.049	0.018	0.314	0.256	0.251
0.1750	0.470	0.988	0.018	0.019	0.312	0.251	0.249
0.2000	0.466	0.996	0.005	0.021	0.310	0.247	0.246
Panel B: progressivity, at $\bar{T}/\bar{Y} = \bar{T}/\bar{Y}^{US}$							
τ_p	$G(h)$	$G(e)$	$G(e_{\text{tot}})$	$G(l)$	$G(y)$	$G(\tilde{y})$	$G(c)$
0.0000	0.493	0.777	0.293	0.033	0.333	0.320	0.298
0.0500	0.492	0.784	0.279	0.029	0.331	0.303	0.282
0.1000	0.491	0.792	0.263	0.026	0.329	0.287	0.266
0.1370	0.490	0.808	0.248	0.027	0.329	0.275	0.255
0.1990	0.487	0.835	0.217	0.022	0.326	0.253	0.236
0.2500	0.484	0.864	0.182	0.020	0.322	0.235	0.221
0.3000	0.481	0.890	0.147	0.017	0.319	0.217	0.206

Notes: $e_{\text{tot}} \equiv e^* + \nu\bar{E}$ for private-school households and $\bar{E}$ for public-school households. Income and consumption Ginis are ε-marginalized, matching the OECD construction. Source: author's calculations.

First, the Gini of private bequests $G(e)$ rises steeply with fiscal size, because an increasing mass of households exits the private-bequest pool entirely. That overstates inequality in education *received*: counting public provision as part of every household's schooling input, the Gini of total education $G(e_{\text{tot}})$ rises only while the public option is still small, peaks near $\bar{T}/\bar{Y} = 0.075$, and then *falls* steeply, from 0.330 at the peak to 0.005 at the top of the grid. Between the two calibrated fiscal sizes it falls from 0.248 to 0.155 along this panel, which holds progressivity at the U.S. value; at the two calibrated regimes proper, which also differ in τ_p, the fall is from 0.248 to 0.119. Either way the ordering on education inputs is the opposite of the ordering on private spending.

Second, the two instruments act asymmetrically on disposable-income inequality. Raising the size of government at fixed progressivity moves $G(\tilde{y})$ comparatively little; raising progressivity at fixed size lowers it monotonically and substantially. Compression of post-tax dispersion in this model is a progressivity phenomenon rather than a fiscal-scale one. That is why a one-parameter flat tax cannot represent the cross-Atlantic pattern, and why both instruments are needed.

Third, the levels in Table C.1 understate measured inequality, but part of that is a reporting convention rather than a model failure. These Ginis marginalize the earnings shock out; the annual survey income underlying the OECD series does not. Counting the shock, the convention used in Table 5 and the one consistent with placing it in the denominator of ι_l, raises the U.S. pre-tax Gini from 0.329 to 0.422 and the disposable Gini from 0.275 to 0.354, against data of 0.517 and 0.395. What remains is concentrated at the top. The model is restricted to labor income of working-age households, abstracts from capital income, transfers outside the modeled instrument, retirees and non-workers, and has no Pareto upper tail. Appendix D.1 adds non-labor income and reports how far that closes what is left.

C.3 Luck and foreseeable variation at the calibrated regimes

The luck share of equation (14) is also ∪-shaped in fiscal size, and for a related reason. As the public-school margin opens, households sort across it, which *polarizes* the distribution of human capital and raises the dispersion of foreseeable log income; once public provision becomes close to universal the distribution compresses again. At the calibrated regimes the model delivers $\iota_l^{\mathrm{US}} = 0.423$ and $\iota_l^{\mathrm{EU}} = 0.438$, a cross-Atlantic gap of +1.5 percentage points.

The nested notions of Section 4.2.4 turn out to be almost indistinguishable here. Adding the inborn-ability term of equation (C.1) moves the share by about two hundredths of a percentage point at either regime, because ξ affects current income only through its influence on the labor-supply decision, and that influence is small, with the Gini of hours below 0.04 at both regimes. Inborn ability in this model operates on the *next* generation's human capital, not on the current cross-section of earnings. The third notion counts inherited human capital as circumstance as well, and it equals one identically. That is a statement about its numerator, which then contains the whole variance of $\log y$: with ε, ξ and h all classified as circumstance, and with labor supply and education spending deterministic functions of (h,ξ), no foreseeable residual is left to sit in the denominator alone. Classifying ability as luck therefore has almost no effect on the decomposition, while classifying inherited human capital as circumstance exhausts it. That is a property of the accounting identity and of an environment in which labor supply and education spending are deterministic functions of the state, not a conclusion about desert, which this model has no apparatus to reach.

Section 4.4.4 reports that the unexplained share of the cross-sectional variance of log earnings in the PSID is 0.705 (s.e. 0.006), falling to 0.287 (s.e. 0.004) once the person-fixed component, which the household knows, is removed. The model's ι_l of 0.42 lies below the econometrician's unexplained share of 0.705, as it should, since households act on determinants of earnings the first-stage regression does not contain. It exceeds the tighter bound of 0.287 by a wide margin, and that overshoot is a real tension rather than a pass. Its likely source is the benchmark restriction $\kappa = 0$: with inborn ability i.i.d. across generations, persistent family-level variation that the data assign to a person-fixed effect has nowhere to go in the model except into ε. Consistent with that reading, the $\kappa = 0.4$ variant in Appendix D.4 lowers the luck share. The level of ι_l should be read as an order of magnitude that the data bound loosely from above, not as a moment the model has matched.

Belief data are not used as a test here, for three reasons. The model's ι_l is a variance ratio and the World Values Survey moment of Section 4.1 is a share of respondents, so no ratio of the two has an interpretation. Beliefs about effort and reward in the United States also predate the modern fiscal system, and may cause it as much as reflect it (Alesina and Angeletos, 2005; Bénabou and Tirole, 2006), a possibility an environment with no belief formation cannot address. And a respondent asked whether success reflects luck may be weighing inheritance, connections, health, or family background, none of which is the ε of equation (10). The belief gap motivates this class of model; it is not a moment this one can be scored on.

C.4 Reclassifying ability and inheritance as luck

Which variation counts as luck is a modeling choice, and the model is transparent about it. Equation (14) counts only ε. But ξ is itself drawn, and h is inherited; on most moral intuitions the family one is born into is a paradigmatic piece of luck (Roemer, 1998). Decomposing $\mathrm{Var}_\mu(\log y^{\mathrm{det}})$ by the law of total variance conditional on h,

$$\mathrm{Var}_\mu\Big(\log y^{\mathrm{det}}\Big) = \underbrace{\mathrm{Var}_h\Big(\mathbb{E}_\xi[\log y^{\mathrm{det}} \mid h]\Big)}_{\text{inherited human capital}} + \underbrace{\mathbb{E}_h\Big(\mathrm{Var}_\xi[\log y^{\mathrm{det}} \mid h]\Big)}_{\text{inborn ability, via } l^*}, \tag{C.1}$$

gives a nested family: $\iota^{(1)}$ counts ε alone; $\iota^{(2)}$ adds the inborn-ability term; $\iota^{(3)}$ adds inherited human capital as well. The third is identically one. That is a property of the environment rather than a defect of the accounting. Once inborn ability and inherited human capital are both classified as circumstance, nothing remains, because labor supply and education spending are deterministic functions of (h, ξ). A model in which $\iota^{(3)}$ were interior would need a source of effort that is orthogonal to circumstance, which this environment, like most of the quantitative literature it builds on, does not contain. Appendix C.3 reports $\iota^{(1)}$ and $\iota^{(2)}$ and shows that they are numerically almost identical here, because ξ operates on the *next* generation's human capital rather than on the current cross-section of income.

C.5 Atkinson aggregation of the welfare results

Table C.2: Atkinson-aggregated consumption-equivalent welfare

Counterfactual	$\varepsilon_A = 0$	0.5	1	2
U.S. → E.U. one-shot	+0.41%	+0.41%	+0.40%	+0.40%
U.S. → E.U. phased (5 gen.)	+0.13%	+0.13%	+0.13%	+0.12%
E.U. → U.S. one-shot	−0.49%	−0.49%	−0.49%	−0.50%
E.U. → U.S. phased (5 gen.)	−0.31%	−0.31%	−0.31%	−0.31%

Notes: The index is applied to the consumption ratios $1+\zeta$, which are strictly positive, and the reported figure is the aggregate less one; the aggregator is therefore well defined for reforms whose average effect is negative. $\varepsilon_A = 0$ gives the arithmetic mean of those ratios and $\varepsilon_A \to 1$ the geometric mean, which is the ζ_{agg} of Table 6. Source: author's calculations.

D Economic Sensitivities and Extensions

D.1 Non-labor income

The environment has no savings margin, so households cannot self-insure against ε and all cross-country differences in insurance load onto the tax-and-transfer system. Following Birinci et al. (2026), I take the intermediate route of endowing households with non-labor income calibrated to its empirical cross-sectional distribution, via the schedule in equation (11).

The calibration comes from the PSID-SHELF panel (Daumler et al., 2025), pooling 1984–2019 waves.[2] Non-labor income is 17.6% of permanent household income in aggregate, and its elasticity with respect to permanent family labor earnings is $\chi = 0.212$ (s.e. 0.028). Because the elasticity is well below one, the non-labor *share* of a household's own income falls steeply across the distribution, from 54% in the bottom decile of permanent labor earnings to 13% in the top. The composition matters here, because the omitted resources are dominated by transfers, which are concentrated at the bottom, rather than by asset income, which is concentrated at the top. Imputing asset income at four percent of net worth excluding housing gives an aggregate share of 8.2% and a level rising by roughly a factor of ten from the bottom decile to the top. It is top-heavy, as expected, but smaller than the transfer component it is bundled with.

Table D.1: Adding exogenous non-labor income

	Benchmark		With non-labor income	
Moment	U.S.	Europe	U.S.	Europe
Hours L	0.268	0.250	0.278	0.262
Population in public school	0.772	0.907	0.719	0.888
Pre-tax Gini $G(y)$	0.329	0.319	0.343	0.330
Disposable Gini $G(\tilde{y})$	0.275	0.245	0.248	0.221
Luck share ι_l	0.423	0.438	0.400	0.419
ζ_{agg} (%)	+0.40 / -0.49		-2.50 / +2.26	
Support share (%)	+75.73 / +22.02		+0.00 / +100.00	
D1−D10 spread (pp)	+0.76 / -0.69		+0.59 / -0.58	

Notes: Non-labor income follows equation (11) with aggregate share $\omega_b = 0.176$ and elasticity $\chi = 0.212$, both estimated on PSID-SHELF. The right-hand block re-calibrates (μ_ξ, ϕ) to the same laissez-faire targets. Welfare rows report the U.S.-to-Europe and Europe-to-U.S. one-shot counterfactuals. Source: author's calculations.

Table D.1 reports the result. Re-calibrating to the same laissez-faire targets with non-labor income switched on leaves the cross-Atlantic signs intact and moves magnitudes. Cross-decile welfare dispersion falls, as one would expect once the poorest households have a resource floor that does not depend on their own earnings. It confirms by computation what Section 4.6.2 could only assert, namely that the dispersion reported there is an

[2] Reference persons aged 25–60 with positive family income and positive family labor earnings; non-labor income is total family income net of the labor earnings of the reference couple, so it comprises capital income, private transfers, and public transfers outside the modeled instrument. Permanent measures are person means over three or more observations: 15,116 persons.

upper bound. The aggregate welfare effects move in the opposite direction and become larger in absolute value. The extension reverses the sign of the benchmark welfare effect, from +0.40% to −2.50%. The sign also remains sensitive to where the anchors are placed, as Appendix D.2 shows, but the sharper point is that the benchmark number does not survive the addition of a resource the data measure and the model otherwise omits.

D.2 Placement of the fiscal anchors

Section 5 settles the numerical question. An economic one remains. The welfare figure depends on where the calibrated economies sit relative to the turning point of the aggregate response, and the calibration places them there using a single moment. Table D.2 scales both fiscal anchors by a common factor s, holding progressivity and everything else at the benchmark, and reports what the model then says.

Table D.2: What moves, and what does not, as the fiscal anchors are rescaled

	Anchors		Cross-Atlantic gaps (Europe − U.S.)				U.S.→Europe	
s	$\bar{T}/\bar{Y}^{\mathrm{US}}$	$\bar{T}/\bar{Y}^{\mathrm{EU}}$	ΔY	ΔL	$\Delta G(\tilde{y})$	$\Delta\iota_l$	ζ_{agg}	Support
0.50	0.0487	0.0583	−0.237	−0.019	−0.017	−0.3	−1.34%	11%
0.66	0.0643	0.0770	−0.245	−0.019	−0.022	+0.7	−2.23%	0%
0.75	0.0730	0.0874	−0.208	−0.020	−0.025	+1.1	−1.86%	1%
0.85	0.0828	0.0991	−0.169	−0.019	−0.028	+1.4	−1.20%	6%
1.00*	0.0974	0.1166	−0.094	−0.018	−0.029	+1.5	+0.40%	76%
1.15	0.1120	0.1341	−0.010	−0.017	−0.029	+1.3	+1.91%	97%
1.30	0.1266	0.1516	+0.074	−0.016	−0.027	+0.9	+3.87%	100%
Data	—	—	< 0	−0.039	−0.099	—	—	—

Notes: Both fiscal anchors are multiplied by s, holding progressivity and every other parameter at the benchmark calibration; $s = 1$ (marked *) is the benchmark, at which the model reproduces each region's public share of education expenditure. $\Delta\iota_l$ is in percentage points. The final two columns report the aggregate consumption-equivalent welfare effect of the one-shot U.S.-to-Europe reform and the share of the initial cross-section preferring it. The data row gives the sign or value of the corresponding empirical gap. Source: author's calculations.

Two readings are worth separating. The placement *is* identified. The public financing share of education rises steeply in s and selects $s \approx 1$, and at that value the output ranking is also correct, so the two informative moments agree rather than conflict. What the table shows is that the welfare figure is nonetheless sensitive to placement, running from −2.2% to +3.9% over the range, and that the moments this literature more commonly matches would not detect the difference. The hours gap varies by less than half a percentage point across the whole range and the disposable-income Gini gap stays correctly signed throughout. A calibration disciplined on hours and inequality alone, without the education-financing target, would leave the welfare sign undetermined. That is an argument for reporting which moment does the identifying work, not a claim that the model fails to identify it.

D.3 The productivity of public spending

The benchmark assumes a dollar of public education spending buys as much human capital as a dollar of private spending, following Bénabou (2002) and Torul (2020). Table D.3 relaxes that. A unit of public spending delivers $\theta \le 1$ units of educational input, and at each θ both fiscal anchors are re-calibrated to reproduce the observed education-financing shares. The laissez-faire calibration is unaffected, since public provision is zero at $\bar{T}/\bar{Y} = 0$.

Table D.3: A public-provision efficiency wedge, with the fiscal anchors re-calibrated at each value

	Re-calibrated anchors		Education share		Output		U.S.→Europe	
θ	$\bar{T}/\bar{Y}^{\mathrm{US}}$	$\bar{T}/\bar{Y}^{\mathrm{EU}}$	U.S.	Europe	ΔY	sign	ζ_{agg}	Support
1.00*	0.0977	0.1163	0.698	0.869	−0.096	yes	+0.33%	73%
0.80	0.1189	0.1415	0.696	0.869	−0.099	yes	−0.14%	48%
0.65	0.1415	0.1680	0.693	0.870	−0.102	yes	−0.62%	23%
0.50	0.1773	0.2066	0.699	0.868	−0.109	yes	−1.65%	1%
0.35	0.2345	0.2703	0.697	0.868	−0.113	yes	−2.85%	0%
0.25	0.2995	0.3394	0.697	0.872	−0.106	yes	−4.17%	0%
Data	—	—	0.696	0.870	< 0	—	—	—

Notes: θ is the educational input delivered by one unit of public spending, against one unit for private spending; $\theta = 1$ (marked *) is the benchmark of Sections 4.5–5. The $\theta = 1$ row re-runs the anchor calibration from scratch rather than importing the benchmark, so it reproduces it only to within the bisection tolerance. At each θ both fiscal anchors are re-calibrated to reproduce the two observed education-financing shares. The laissez-faire calibration of (μ_ξ, ϕ) is unaffected by θ, since public provision is zero at $\bar{T}/\bar{Y} = 0$. Source: author's calculations.

The output ranking is correct at every value, including the benchmark $\theta = 1$, so the wedge is not needed to reconcile the model with the data, which is itself a change from what the coarse solution implied. What it does affect is the welfare figure, which falls monotonically in the wedge and turns negative below roughly $\theta = 0.9$. A reader who believes public provision is materially less productive per dollar than private provision should read the European reform as welfare-reducing rather than mildly welfare-improving. The parameter is not disciplined here, and I do not claim a value for it; the point is to record how much of the welfare conclusion rests on an assumption the antecedent literature makes implicitly. The lower rows do carry their own warning, though. At $\theta = 0.25$ the calibration requires an instrument of 30% of aggregate labor income devoted to schooling and cash transfers alone, before pensions, health care or social insurance, which no observed fiscal system resembles. Whatever the true wedge is, the rows below roughly $\theta = 0.5$ are arithmetically consistent and economically implausible, and the welfare figures there should be read as showing the direction of the effect rather than its size.

D.4 Parametric robustness

Table D.4: Parametric robustness at the calibrated regimes

	$\bar{H}$		L		ι_l gap
Specification	U.S.	Europe	U.S.	Europe	(pp)
canonical	0.320	0.279	0.268	0.250	+1.50
κ =0.4 (AR(1) ability)	0.644	0.521	0.267	0.249	+1.97
ν =0 (no spillover)	0.219	0.226	0.266	0.249	+1.23
ν =1 (full spillover)	0.586	0.518	0.272	0.253	+0.59
ψ =0.20	0.373	0.420	0.271	0.255	+0.80
ψ =0.80	0.546	0.431	0.271	0.252	−0.54
ω =0.25 (Frisch 4)	0.247	0.212	0.205	0.189	+1.51
ω =1.00 (Frisch 1)	0.439	0.390	0.372	0.353	+1.41
γ =0.10	0.315	0.306	0.266	0.249	+1.21
γ =0.30	0.334	0.273	0.269	0.250	+1.73
α =0.20	0.861	1.019	0.264	0.248	+0.53
α =0.40	0.290	0.208	0.274	0.253	+1.94
λ =0.55	0.254	0.254	0.300	0.281	+1.62
λ =0.70	0.442	0.354	0.231	0.215	+1.30

Notes: Each row varies one parameter, holding the remainder at the benchmark calibration and the fiscal instruments at their calibrated values. The final column is the cross-Atlantic gap in the luck share of equation (14), in percentage points. Source: author's calculations.

Every row holds the fiscal anchors at their benchmark values rather than re-calibrating them, so the rows isolate the effect of the parameter on the equilibrium at a fixed policy, not the combined effect of parameter and re-anchoring. The cross-Atlantic ordering of hours is preserved in every specification. Two rows bear on Section 5. The education elasticity α governs the location of the turning point, and at $\alpha = 0.20$ the ordering of aggregate human capital reverses, 0.861 in the U.S. against 1.019 in Europe, while at $\alpha = 0.40$ it is preserved and sharpened. That reversal is an artifact of the fixed-policy design, however, and Appendix D.5 shows it disappears once the anchors are re-calibrated at each α. It is the lesson of Section 5 reached from the parameter side rather than the grid side. What survives the re-calibration is that α moves the trough, and with it the welfare number. It is a parameter taken from the literature rather than estimated here, and the paper's welfare conclusion is conditional on it. The luck-share gap is small and positive throughout. Combined with the provenance of σ_ξ^2 set out in Section 4.4.4, its narrowness is a further reason not to read the composition of pre-tax inequality as a sharp prediction of the model. Section 5.2 reports the grid-resolution check and Appendix G the phased-reform horizons.

D.5 The education elasticity

Section 4.5.1 identifies α, the elasticity of human capital to the education input, as the parameter that locates the trough of the aggregate response, and therefore as the economic counterpart of the numerical result. Table D.4 varies it at fixed policy. That is the wrong exercise for this question, because if α moves the trough, the anchors must move with it, since they are calibrated to a moment the trough's location affects. Table D.5 re-bisects both anchors at each value.

Table D.5: The education elasticity, with both fiscal anchors re-calibrated at each value

	Re-calibrated anchors		Education share		Output		U.S.→Europe
α	$\bar{T}/\bar{Y}^{\mathrm{US}}$	$\bar{T}/\bar{Y}^{\mathrm{EU}}$	U.S.	Europe	ΔY	sign	ζ_{agg}
0.20	0.0686	0.0832	0.696	0.870	−0.088	yes	+1.50%
0.25	0.0835	0.1005	0.696	0.870	−0.098	yes	+0.87%
0.30*	0.0973	0.1166	0.696	0.871	−0.095	yes	+0.40%
0.35	0.1102	0.1316	0.695	0.870	−0.095	yes	−0.05%
0.40	0.1229	0.1458	0.697	0.870	−0.102	yes	−0.63%
Data	—	—	0.696	0.870	<0	—	—

Notes: α is the elasticity of next-period human capital to the education input in equation (9); $\alpha = 0.30$ (marked *) is the benchmark, and that row reproduces the benchmark calibration. At every value both fiscal anchors are re-bisected to the two observed education-financing shares, so each row is a complete exercise rather than a parameter moved at fixed policy. The final column is the aggregate consumption-equivalent effect of the one-shot U.S.-to-Europe reform along a $T = 25$ transition. Source: author's calculations.

Three points come out of the table. First, the calibration absorbs α in the direction one would expect: a lower elasticity requires a larger government to deliver the same public financing share, so the anchors rise monotonically from 0.069 to 0.123 across the range. Second, the untargeted output ranking is correct at every value, including $\alpha = 0.20$, where Table D.4 reported the ordering of aggregate human capital reversing at fixed policy. Re-calibrating restores it, which is a further illustration of the paper's theme from the opposite direction, since holding the calibration fixed while moving a parameter can be as misleading as holding it fixed while moving the grid.

Third, and most important for how much weight the welfare number can bear, ζ_{agg} falls monotonically in α and changes sign between 0.30 and 0.35, running from +1.50% to −0.63% across a range that brackets the values used in this literature. That span is an order of magnitude wider than the 0.22 percentage points of residual numerical variation in Table 9. The economic uncertainty about α dominates the numerical uncertainty about the grid, which is the right ordering of this paper's two kinds of sensitivity and the reason Section 8 reports an interval rather than a point estimate for the transatlantic comparison. It does not weaken the methodological result, which is about how discretization reaches a conclusion and not about how large this particular conclusion is.

D.6 Sensitivity to the progressivity anchor

The benchmark takes τ_p from Holter et al. (2019). Using instead the Heathcote et al. (2017) U.S. estimate of 0.181 raises U.S. progressivity toward the European value and therefore compresses every cross-Atlantic gap that operates through the progressivity channel, without changing signs. As Section 4.4.3 explains, the two estimates are taken on different income concepts, and the Holter et al. concept is the one the model's tax base corresponds to.

The reader should weigh these sensitivities against the numerical ones. The welfare figure moves by 0.22 percentage points across fourteen discretizations (Table 9) and by several percentage points across the economic variations above. The economic uncertainty dominates, which is the right ordering, and it is why the substantive transatlantic claim in this paper is an interval rather than a point.

E Policy preferences under majority rule

This appendix computes the pairwise majority tournament over the $(\bar{T}/\bar{Y}, \tau_p)$ grid for each household in the status-quo cross-section, using the transition-based valuations of Section 4. Table E.1 reports it under both value concepts and both cross-sections.

Because Stiglitz (1974) showed that preferences over taxes are generically double-peaked when public and private provision coexist, a property Torul (2020) documents in this environment, the median-voter theorem cannot be invoked. Following Carroll et al. (2021) I compute the full pairwise majority tournament over the policy grid, evaluating each policy along a one-shot perfect-foresight transition from the status quo rather than by comparing stationary equilibria, and search for a Condorcet winner.

Table E.1: Majority rule over the fiscal policy grid

Value concept and cross-section	Condorcet winner $(\bar{T}/\bar{Y}, \tau_p)$	Uncovered set size	Mass at highest $\bar{T}/\bar{Y}$	Mass at interior τ_p
Transition value, μ^{US}	(0.20, 0.30)	1	1.000	0.316
Transition value, μ^{EU}	(0.20, 0.30)	1	1.000	0.251
Steady-state value, μ^{US}	(0.20, 0.14)	1	1.000	0.684
Steady-state value, μ^{EU}	(0.20, 0.14)	1	1.000	0.688

Notes: Pairwise majority tournament over the 63 policies of the $(\bar{T}/\bar{Y}, \tau_p)$ grid. A Condorcet winner beats every other policy in a pairwise majority comparison; the uncovered set is that of McKelvey (1986). "Transition value" evaluates each policy along a one-shot perfect-foresight transition from the status quo; "steady-state value" compares stationary equilibria. The two agree on the size of government and disagree on progressivity. Source: author's calculations.

A Condorcet winner exists in every case and the uncovered set is a singleton, but the result is degenerate rather than informative. The mass at the largest fiscal size is 1.000 throughout, so unanimity delivers the winner and rules out cycling without majority rule doing any work. The two value concepts agree on fiscal scale and disagree on progressivity, which is the substantive content of the exercise. The tournament is also bounded by the policy grid. Because the winner falls at its upper edge, the model identifies a direction rather than an equilibrium tax rate. It also reads each household's preference entirely off (h, ξ), whereas Carroll et al. (2025) find survey-measured attitudes toward redistribution tracking political identity more strongly than economic circumstance. The model contains no parties, agenda setting, or lobbying, and the exercise is not a positive theory of how fiscal policy is chosen.

F Levels versus logs in the variance decomposition

Section 4.2.4 takes the decomposition of pre-tax income variation in logs. An alternative, used in the antecedent literature, writes income as $y = \Theta l^{1-\lambda} h^{\lambda} + \eta$ with an additive shock and decomposes the variance of y in levels. That specification has three drawbacks, none of them presentational.

The share is not invariant to units: With an additive shock of fixed variance σ_η^2, the foreseeable component $\mathrm{Var}_\mu(\Theta l^{1-\lambda} h^{\lambda})$ scales with the square of the income unit while σ_η^2 does not. Multiplying the TFP normalization Θ by two, a change of units with no economic content, moves the implied luck share at the U.S. regime from 0.379 to 0.041, a factor of more than nine. The same operation under the log specification of equation (14) moves it from 0.423 to 0.435, a factor of 1.03, and that residual movement is a genuine equilibrium response rather than a normalization artifact, since Θ is not a pure scaling in an environment where education spending and consumption share units. Each specification is evaluated at its own calibrated (μ_ξ, ϕ), so the comparison is between two internally consistent economies.

The calibrated shock variance is not the object the source estimates: The residual variance of log wages reported by Heathcote et al. (2010) is the variance of a log residual. Inserting it as the variance of an additive shock to the *level* of income implies a shock whose standard deviation is 43% of mean income, which is not what was estimated.

Negative income realizations force an artificial floor: Under the additive specification the lower tail of the shock drives pre-tax income below zero for low human-capital households, so consumption must be bounded below by an arbitrary floor. In the calibrated version of that specification the floor binds on 14.4% of the probability mass in the bottom decile of the human-capital distribution, which is where the valuations driving both the cross-decile welfare statistics and the distribution of policy preferences come from, and on 1.6% of the cross-section overall. Under equation (5) income is strictly positive and no floor is required; the numerical guard in the code does not bind at either calibrated regime.

The substantive consequence is that the two specifications disagree about the model's comparative statics. Measured in levels, the foreseeable variance falls between the U.S. and European regimes largely because mean income falls, so the luck share appears to rise with redistribution. Measured in logs, the dispersion of foreseeable income does not fall in the same way, and the luck share is ∪-shaped in fiscal size rather than monotone. The cross-Atlantic gap in the luck share is small and, as Table D.4 shows, not robust in sign.

G Phased reform horizons

The counterfactuals in Section 4.6.2 implement the policy change at a single date. This appendix reports the same reforms phased in linearly over two and five model periods. Each model period is one generation, so the one-period case is the one-shot counterfactual and the five-period case is a long-horizon smoothing benchmark rather than a legislative timetable. Table G.1 reports all three horizons in both directions.

Table G.1: Welfare and dispersion under three phase-in horizons

Direction and phasing	ζ_{agg}	ζ_{med}	Support share	D1−D10 (pp)
U.S. → E.U. reform				
one-period (25 yr)	+0.40%	+0.51%	75.7%	+0.76
two-period (50 yr)	+0.31%	+0.39%	78.5%	+0.46
five-period (125 yr)	+0.13%	+0.17%	77.1%	+0.20
E.U. → U.S. reform				
one-period (25 yr)	−0.49%	−0.60%	22.0%	−0.69
two-period (50 yr)	−0.42%	−0.48%	15.8%	−0.39
five-period (125 yr)	−0.31%	−0.33%	9.2%	−0.15

Notes: One-period phasing is the one-shot counterfactual of Table 6. Each model period is one generation (≈ 25 years). Source: author's calculations.

Phasing reduces the cross-decile dispersion of welfare consequences substantially in both directions, because the dynastic discount factor $\rho = 0.80$ concentrates the welfare integral on early generations and a phased reform leaves those generations close to the status quo. The aggregate welfare effect is less sensitive to the horizon than the dispersion is, and its sign is unchanged throughout.

References

Alesina, A. and Angeletos, G.-M. (2005). Fairness and redistribution. *American Economic Review*, 95(4):960–980.

Bénabou, R. (2002). Tax and education policy in a heterogeneous-agent economy: What levels of redistribution maximize growth and efficiency? *Econometrica*, 70(2):481–517.

Bénabou, R. and Tirole, J. (2006). Belief in a just world and redistributive politics. *Quarterly Journal of Economics*, 121(2):699–746.

Birinci, S., Karabarbounis, L., and See, K. (2026). Why do americans no longer work so much more than non-americans? Working paper, Federal Reserve Bank of St. Louis.

Carroll, C. D. (2006). The method of endogenous gridpoints for solving dynamic stochastic optimization problems. *Economics Letters*, 91(3):312–320.

Carroll, D. R., Dolmas, J., and Young, E. R. (2021). The politics of flat taxes. *Review of Economic Dynamics*, 39:174–201.

Carroll, D. R., Luduvice, A. V. D., and Young, E. R. (2025). A note on aggregating preferences for redistribution. *Economics Letters*, 256:112535.

Daumler, D., Friedman, E., and Pfeffer, F. T. (2025). Psid-shelf user guide and codebook, 1968–2021. Institute for Social Research, University of Michigan; ICPSR Open Data Release 2025-01 (DOI:10.7302/25205). Panel Study of Income Dynamics–Social, Health, and Economic Longitudinal File (PSID-SHELF), 1968–2021 beta.

Heathcote, J., Perri, F., and Violante, G. L. (2010). Unequal we stand: An empirical analysis of economic inequality in the united states, 1967–2006. *Review of Economic Dynamics*, 13(1):15–51.

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

Holter, H. A., Krueger, D., and Stepanchuk, S. (2019). How do tax progressivity and household heterogeneity affect Laffer curves? *Quantitative Economics*, 10(4):1317–1356.

Kopecky, K. A. and Suen, R. M. H. (2010). Finite state Markov-chain approximations to highly persistent processes. *Review of Economic Dynamics*, 13(3):701–714.

Krueger, D., Perri, F., Pistaferri, L., and Violante, G. L. (2010). Cross-sectional facts for macroeconomists. *Review of Economic Dynamics*, 13(1):1–14.

MacQueen, J. and Porteus, E. L. (1975). Numerical methods and bounds in markovian decision processes. *Journal of Mathematical Analysis and Applications*, 49(2):421–442.

McKelvey, R. D. (1986). Covering, dominance, and institution-free properties of social choice. *American Journal of Political Science*, 30(2):283–314.

Roemer, J. E. (1998). *Equality of Opportunity*. Harvard University Press, Cambridge, MA.

Stiglitz, J. E. (1974). The demand for education in public and private school systems. *Journal of Public Economics*, 3(4):349–385.

Tauchen, G. (1986). Finite state markov-chain approximations to univariate and vector autoregressions. *Economics Letters*, 20(2):177–181.

Torul, O. (2020). Education in a heterogeneous-agent economy: Revisiting transatlantic differences. *Journal of Human Capital*, 14(2):165–216.