Appendix A: The Data Sources

We have used the following data sources in this study:

(a) **HMRC (2015)**. A disaggregation of HMRC tax receipts between England, Wales, Scotland & Northern Ireland: Methodology Note.
(b) **HMRC (2014)**. Income tax liabilities statistics: number of individual income taxpayers by region (Table 2.2) and share of total income for percentile groups (Table 2.4).
(c) **Department for Communities and Local Government (2015)**. Band D council tax for LA's.
(d) **Department for Communities and Local Government (2011)**. Number of all chargeable dwellings.
(e) **ONS (2016)**. Geometric centroid for each LA from GIS.
(f) **ONS (2015)**. Regional accounts: gross value added (GVA) measure, Welsh economic region and year.
(g) **ONS (2014)**.House price statistics for small areas: median sale price by dwelling type and LA.
(h) **ONS (2014)**. Labour market statistics: population, employment, unemployment, inactivity and job density.
(i) **ONS (2015, 2014)**. Migration statistics unit: internal migration between English and Welsh LAs.
(j) **ONS (2012)**.Small area income estimates: total household weekly income.
(k) **ONS (2011)**. Squared Euclidean distance matrix at LA level.
(l) **Stats Wales (2015)**.Number of all chargeable dwellings.
(m) **Welsh Government (2015)**.Council tax levels by billing authority and band.

The data are used for two purposes: (A) calibration of the economic model, or/and (B) estimation of the econometric model. There are three ways of using the data: first, direct use as one of the regressors in the econometric model; second, calibration of the parameters in the economic model by matching the model-implied endogenous variables with the observed endogenous variables; third, as the basis to implement an Extended Monte Carlo (EMC) simulation procedure (detailed in Appendix C) to generate the data needed for the econometric or economic models. The variables used in the analysis, the data sources and the techniques are summarised in the following table.

Variable	Source	Purpose	Technique
Council tax distribution	(d) (l)	(B)	EMC
Council tax rates	(c) (m)	(A) (B)	Derived
Euclidean distance	(k)	(B)	Direct use
Geographic distance	(e)	(B)	Pythagoras theorem
Government spending	(a)	(A)	Calibration
House price	(g)	(B)	EMC
Inactivity	(h)	(A) (B)	Calibration, Direct use
Income tax distribution	(b)	(A) (B)	Calibration, EMC
Job density	(h)	(B)	Direct use
Migration flow	(i)	(A) (B)	Direct use
Output	(f)	(A)	Calibration
Population	(h)	(B)	Direct use
Tax revenue	(a)	(A)	Calibration
Unemployment	(h)	(A) (B)	Calibration, Direct use
Wage	(j)	(A) (B)	EMC

Appendix B: The CGE Model of the Devolved Economy

The Consumer

There are three types ($i = 1, 2, 3$) of consumers/taxpayers and wage (w_i for pre-tax wage, ω_i for after-tax wage). Everyone receives a lump-sum benefit (bf) from the government common to all types. Utility depends on the private goods (z_i), composed of housing (h_i) and other private consumption (c_i), and public goods provided by the local governments (G) such as education and healthcare. Note that the income redistribution, such as jobseeker's allowance $\underline{\omega}$ for the unemployed (u is the unemployment rate relevant only to type 1 low-income consumers), is paid directly by the central government, so it can be treated as exogenous. There are different types of tax rates (τs) under different income thresholds (κs), which are summarised in Table B1 (as of 2014).

Thresholds	**Income Bands** ω_i		**Employee Rates**		**Employer Rates**
κ	Lower	Upper	τ_{IT}	τ_{NI}	τ_{NIF}
	0	8,064	0%	0%	0%
κ_{NI1}	8,064	8,112	0%	12%	0%
κ_{NIF}	8,112	10,600	0%	12%	13.80%
κ_{IT1}	10,600	42,385	20%	12%	13.80%
$\kappa_{NI2} = \kappa_{IT2}$	42,385	150,000	40%	2%	13.80%
κ_{IT3}	150,000	∞	45%	2%	13.80%

Table B1 Combined Bands for Income Tax Rates and NI Rates

The maximisation problem of a low-income (type 1) consumer is:

$$\max_{z_1, c_1, h_1} u_1(z_1, G) = \left[\alpha_1 z_1^{\frac{s_1 - 1}{s_1}} + (1 - \alpha_1) G^{\frac{s_1 - 1}{s_1}} \right]^{\frac{s_1}{s_1 - 1}}, \text{ subject to:}$$

(H1.1) Budget Constraint: $(1-u)\omega_1 + u\underline{\omega} + bf = p(1+\tau_C)c_1 + p_{H1}(1+\tau_{H1})h_1$

(H1.2) Income: $\omega_1 = \kappa_{NI1} + (\kappa_{IT1} - \kappa_{NI1})(1-\tau_{NI1}) + (w_1 - \kappa_{IT1})(1-\tau_{NI1} - \tau_{IT1})$

(H1.3) Private Goods Nesting: $z_1 = \left[\beta_1 c_1^{\frac{ss_1 - 1}{ss_1}} + (1-\beta_1) h_1^{\frac{ss_1 - 1}{ss_1}} \right]^{\frac{ss_1}{ss_1 - 1}}$.

Similarly, the maximisation problem of a middle-income (type 2) consumer is:

$$\max_{z_2, c_2, h_2} u_2(z_2, G) = \left[\alpha_2 z_2^{\frac{s_2 - 1}{s_2}} + (1 - \alpha_2) G^{\frac{s_2 - 1}{s_2}} \right]^{\frac{s_2}{s_2 - 1}}, \text{ subject to:}$$

(H2.1) Budget Constraint: $\omega_2 + bf = p(1+\tau_C)c_2 + p_{H2}(1+\tau_{H2})h_2$

(H2.2) Income: $$\omega_2 = \kappa_{NI1} + (\kappa_{IT1} - \kappa_{NI1})(1-\tau_{NI1}) + (\kappa_{IT2} - \kappa_{IT1})(1-\tau_{NI1} - \tau_{IT1}) + (w_2 - \kappa_{IT2})(1-\tau_{NI2} - \tau_{IT2})$$

(H2.3) Private Goods Nesting: $$z_2 = \left[\beta_2 c_2^{\frac{ss_2-1}{ss_2}} + (1-\beta_2) h_2^{\frac{ss_2-1}{ss_2}} \right]^{\frac{ss_2}{ss_2-1}}.$$

And the maximisation problem of a top-income (type 3) consumer is:

$$\max_{z_3, c_3, h_3} u_3(z_3, G) = \left[\alpha_3 z_3^{\frac{s_3-1}{s_3}} + (1-\alpha_3) G^{\frac{s_3-1}{s_3}} \right]^{\frac{s_3}{s_3-1}}, \text{ subject to:}$$

(H3.1) Budget Constraint: $\omega_3 + bf = p(1+\tau_C)c_3 + p_{H3}(1+\tau_{H3})h_3$

(H3.2) Income: $$\omega_3 = \kappa_{NI1} + (\kappa_{IT1} - \kappa_{NI1})(1-\tau_{NI1}) + (\kappa_{IT2} - \kappa_{IT1})(1-\tau_{NI1} - \tau_{IT1}) + (\kappa_{IT3} - \kappa_{IT2})(1-\tau_{NI2} - \tau_{IT2}) + (w_3 - \kappa_{IT3})(1-\tau_{NI2} - \tau_{IT3})$$

(H3.3) Private Goods Nesting: $$z_3 = \left[\beta_3 c_3^{\frac{ss_3-1}{ss_3}} + (1-\beta_3) h_3^{\frac{ss_3-1}{ss_3}} \right]^{\frac{ss_3}{ss_3-1}}.$$

The Producer

There are three types of labour input L_i corresponding with the three types of taxpayers, total factor productivity is A, and p the output price. The representative producer's maximisation problem is:

$$\begin{aligned} \max \Pi = pY &- \left[\kappa_{NIF} + (1+\tau_{NIF})(w_1 - \kappa_{NIF}) \right] L_1 (1-u) \\ &- \left[\kappa_{NIF} + (1+\tau_{NIF})(w_2 - \kappa_{NIF}) \right] L_2 \\ &- \left[\kappa_{NIF} + (1+\tau_{NIF})(w_3 - \kappa_{NIF}) \right] L_3 \end{aligned} \text{, subject to:}$$

(F1) Production Function: $$Y = A \left\{ \gamma_1 \left[(1-u) L_1 \right]^{\frac{\sigma-1}{\sigma}} + \gamma_2 (L_2)^{\frac{\sigma-1}{\sigma}} + \gamma_3 (L_3)^{\frac{\sigma-1}{\sigma}} \right\}^{\frac{\sigma}{\sigma-1}}.$$

The Government

The budget constraint of the Welsh government is[1]:

$$T_H + (T_{ITWales} + B_{UK}) = G$$

The term in the brackets was the original block grant from Whitehall, which did not distinguish between the income tax kept in Wales ($T_{ITWales}$) and the other part (B_{UK}) before the devolution. However, after the 2014 Act, the two items could be explicitly separated to allow for different

[1] Under the Wales Act 2017 the Welsh government can borrow up to £1 billion from central government but this must be used only for capital expenditure which is not considered in this static model.

income tax rates in Wales. In the status quo, the Welsh government is entitled to 10p in the pound of the yield of Welsh income taxes, i.e. the Basic Rate, the Higher Rate and the Additional Rate. For example, the Higher Rate is currently 40p in the pound, so 10p in the pound is kept by Welsh government as part of $T_{ITWales}$. If they decide to raise the Higher Rate in Wales to 41p in the pound, then 11p in the pound will be kept.

(G1) Council Tax Generated and Used in Wales:

$$T_H = p_{H1}\tau_{H1}h_1L_1 + p_{H2}\tau_{H2}h_2L_2 + p_{H3}\tau_{H3}h_3L_3$$

(G2) Income Tax Kept in Wales:

$$\begin{aligned} T_{ITWales} &= \left(w_1-\kappa_{IT1}\right)\left(\tau_{IT1}-10\%\right)L_1\left(1-u\right) \\ &+\left[\left(\kappa_{IT2}-\kappa_{IT1}\right)\left(\tau_{IT1}-10\%\right)+\left(w_2-\kappa_{IT2}\right)\left(\tau_{IT2}-30\%\right)\right]L_2 \\ &+\left[\left(\kappa_{IT2}-\omega_{IT1}^{*}\right)\left(\tau_{IT1}-10\%\right)+\left(\kappa_{IT3}-\kappa_{IT2}\right)\left(\tau_{IT2}-30\%\right)+\left(w_3-\kappa_{IT3}\right)\left(\tau_{IT3}-35\%\right)\right]L_3 \end{aligned}$$

(G3) Income Tax Generated in Wales:

$$\begin{aligned} T_{IT} &= \left(w_1-\kappa_{IT1}\right)\tau_{IT1}L_1\left(1-u\right) \\ &+\left[\left(\kappa_{IT2}-\kappa_{IT1}\right)\tau_{IT1}+\left(w_2-\kappa_{IT2}\right)\tau_{IT2}\right]L_2 \\ &+\left[\left(\kappa_{IT2}-\omega_{IT1}^{*}\right)\tau_{IT1}+\left(\kappa_{IT3}-\kappa_{IT2}\right)\tau_{IT2}+\left(w_3-\kappa_{IT3}\right)\tau_{IT3}\right]L_3 \end{aligned}$$

(G4) National Insurance Contributions Generated in Wales:

$$\begin{aligned} T_{NI} &= \left(w_1-\kappa_{NI1}\right)\tau_{NI1}L_1\left(1-u\right) \\ &+\left[\left(\kappa_{NI2}-\kappa_{NI1}\right)\tau_{NI1}+\left(w_2-\kappa_{NI2}\right)\tau_{NI2}\right]L_2 \\ &+\left[\left(\kappa_{NI2}-\kappa_{NI1}\right)\tau_{NI1}+\left(w_3-\kappa_{NI2}\right)\tau_{NI2}\right]L_3 \\ &+\tau_{NIF}\left[\left(w_1-\kappa_{NIF}\right)L_1\left(1-u\right)+\left(w_2-\kappa_{NIF}\right)L_2+\left(w_3-\kappa_{NIF}\right)L_3\right] \end{aligned}$$

(G5) Consumption VAT Tax Generated in Wales:

$$T_C = p\left(\tau_C c_1 L_1 + \tau_C c_2 L_2 + \tau_C c_3 L_3\right)$$

(G5) Corporation Tax Generated in Wales:

$$T_F = \tau_F \Pi \,.$$

The total tax revenue generated in Wales is $T_{Wales} = T_H + T_{IT} + T_{NI} + T_C + T_F + T_O$, where T_O captures all other tax revenues. It only accounts for a little over 60% of the total government expenditure in Wales (including public goods, pension and welfare benefits).

The remaining expenditure is subsidised by English taxpayers, equivalent to a payment of almost £4000 to each person in Wales. We assume the UK government will fix the block grant, so G can be endogenously derived from the budget constraint.

Model Closure

To solve the model, we need to provide model closure conditions and specify which variables are exogenous. Consumers and firms maximise their objective functions treating the following variables as exogenous: w_i, G, L_i, u.

- Wages (w_i). Imposing the labour markets clearing conditions will determine wages (equivalent to equating wages from the consumers' first order conditions (FOCs) and from the firms' FOCs).
- Government Expenditure on Public Goods (G). The aggregate supply (Y) is demanded for consumption (of the three types of labour force and the pensioners PS), investment (I), or government expenditure on public goods (G). The gap between the supply and demand in Wales is filled by NT[2], which is the net transfer from the rest of the UK. Pensioners' consumption component PS is assumed to be fixed and investment I is assumed to be fixed as a proportion of GDP. Public goods (G) are endogenously determined in the Welsh government's budget constraint. Therefore, imposing the goods market clearing condition will determine the net transfer to Wales (NT) from the rest of the world.
$$Y = \left[(c_1 + h_1) L_1 + (c_2 + h_2) L_2 + (c_3 + h_3) L_3 + PS \right] + I + G(L_1 + L_2 + L_3) + NT \,.$$
- Population (L_i) depends on net immigration (M_i) into Wales,
$$L_i = \overline{L}_i + M_i \,,$$
where $\overline{L}_i$ is the original number of consumer type i and M_i is estimated using the econometric model.
- Unemployment Rate (u). This is exogenous because the economy is modelled in long run equilibrium with unemployment at its equilibrium rate
- Goods Price (p) and House Price (p_{Hi}). Following CGE modelling convention, both are set to 1, so that the quantities can be interpreted as the expenditures[3]. Thus, c_i is interpreted as the total expenditure on consumption and h_i is the total expenditure on housing.

[2] NT is not to be confused with B_{UK}. The latter is transferred to the Welsh government, so it balances the Welsh government's budget constraint. While the former is transferred from the English good market to the Welsh goods market, and it balances aggregate supply and aggregate demand in the goods market in Wales.
[3] Housing (like consumption, investment and G) is a part of Y. There is only one "homogenous" output which can be used for all purposes (including housing). This output is also the numeraire goods for denominating the other price (real wage) in the model.

There are 27 competitive equilibrium conditions for the 30 endogenous variables: $c_1, h_1, \omega_1, z_1, u_1, c_2, h_2, \omega_2, z_2, u_2, c_3, h_3, \omega_3, z_3, u_3, w_1, w_2, w_3, \Pi, Y, I, NT, T_C, T_H, T_{NI}, T_{IT}, T_F, T_{ITWales}$, T_{Wales} and G. There are 12 exogenous variables: $B_{UK}, PS, u, L_1, L_2, L_3, \bar{L}_1, \bar{L}_2, \bar{L}_3, M_1, M_2$ and M_3. Moreover, there are 11 parameters to be calibrated: $\beta_1, \beta_2, \beta_3, ss_1, ss_2, ss_3, \gamma_1, \gamma_2, \gamma_3, \sigma$ and A. Note that other preference parameters, such as $\alpha_1, \alpha_2, \alpha_3, s_1, s_2, s_3$, only exist to define un-observable endogenous variables, so they cannot (and need not to) be estimated based on the data. They can, however, be set at some reasonable values for completeness, but they do not affect the analysis whatsoever. Policy parameters such as τs, κs, bf and $\underline{\omega}$ are all known and set at their actual values.

Using the technique of optimal calibration, the estimated structural parameters governing the behaviour of the consumers and firms are summarised in Table B2:

Parameter	**Meaning**	**Estimate**
β_1	Utility share of consumption (type 1)	0.950
β_2	Utility share of consumption (type 2)	0.276
β_3	Utility share of consumption (type 3)	0.519
ss_1	CES between consumption and housing (type 1)	0.236
ss_2	CES between consumption and housing (type 2)	0.363
ss_3	CES between consumption and housing (type 3)	1.678
γ_1	Income share of type 1 labour/individual	0.639
γ_2	Income share of type 2 labour/individual	0.329
γ_3	Income share of type 3 labour/individual	0.032
σ	CES between the three types of labour	1.056
A	Total factor productivity	79.19

Table B2 Estimated Structural Parameters using Optimal Calibration

Appendix C: Extended Monte Carlo Simulation

The distribution of house price in Figure C1 is based on the 165,822 properties in England and Wales recorded by the Land Registry.

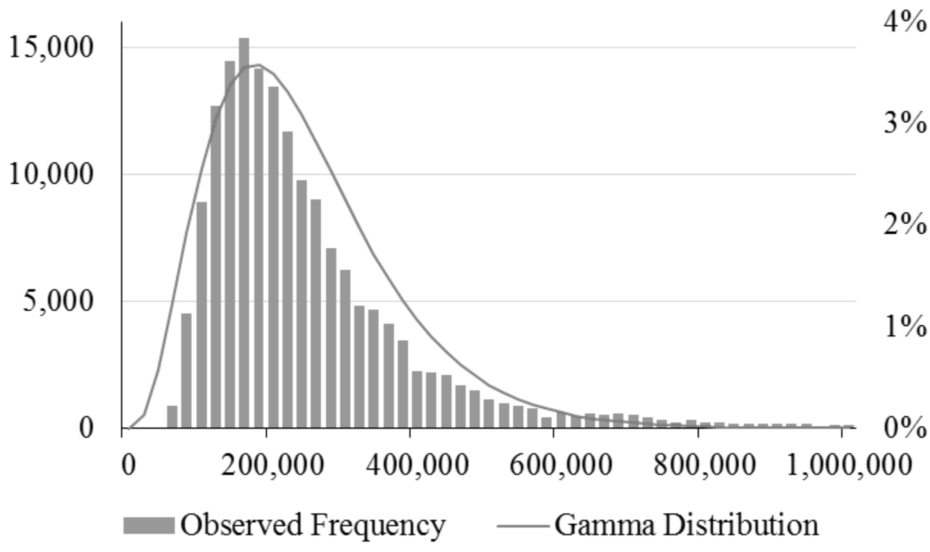

Figure C1 Distribution of House Prices in England and Wales (2010)

Note that we can use a gamma distribution to fit the house price data well, so the information contained in the observed distribution can be summarised by only two parameters parsimoniously. The two parameters (a, b) respectively characterise the shape (a) and the scale (b) of the gamma distribution.

$$gamma(X) = \frac{1}{b^a \Gamma(a)} X^{a-1} \exp\left(-\frac{X}{b}\right)$$

To facilitate the econometric and economic modelling later, we design an Extended Monte Carlo simulation procedure to generate the house prices (p_{H1}, p_{H2}, p_{H3}) for all three types of households.

Step 1: Estimate the shape parameter. Based on the observed numbers of properties across council tax bands at LA level, we can imply the shape of the distribution of house prices – because house prices are strictly increasing with council tax bands. Therefore, the estimated shape parameter $\hat{a}$ of the distribution of council tax bands should be the same as that governing the shape of the gamma distribution of house prices.

Step 2: Derive the scale parameter. Note that the estimated scale parameter $\hat{b}$is not directly applicable to house prices because the horizontal axis of council tax bands are A, B, C, etc. while that of house price is pounds. But we can make use of the observed median house prices in each LA and the relationship among median ($\tilde{p}_H$), mean ($\bar{p}_H$) and the two parameters of gamma distribution to derive the corresponding scale parameter.

$$\left.\begin{aligned} \tilde{p}_H &= \bar{p}_H \frac{3a-0.8}{3a+0.2} \\ \bar{p}_H &= E\left[p_H\right] = ab \end{aligned}\right\} \Rightarrow b = \tilde{p}_H \times \frac{1}{a} \times \frac{3a-0.2}{3a-0.8}$$

Step 3: Simulate the data. In this way, we estimate a unique distribution of house prices in each LA by a parsimonious parametric model (i.e. $\hat{a}$ and $\hat{b}$), based on which, we can simulate $N_S = 10^5$ observations of house prices.

Step 4: Obtain the quantities of interest. With the simulated data, it is easy to obtain the mean/median house prices for the three types of households. For example, we know that the proportion of type 1 household is 89%, then we can use the first 89% simulated house prices (sorted) to calculate the mean/median house price of type 1 household.

The Extended Monte Carlo procedure makes full use of the observed data before standard Monte Carlo simulation, so it has advantages of both bootstrapping and standard Monte Carlo re-sampling techniques. Similarly, this technique is applied to generating mean wages (w_1, w_2, w_3) of each household type in each LA.

Appendix D: Robustness of Conclusions

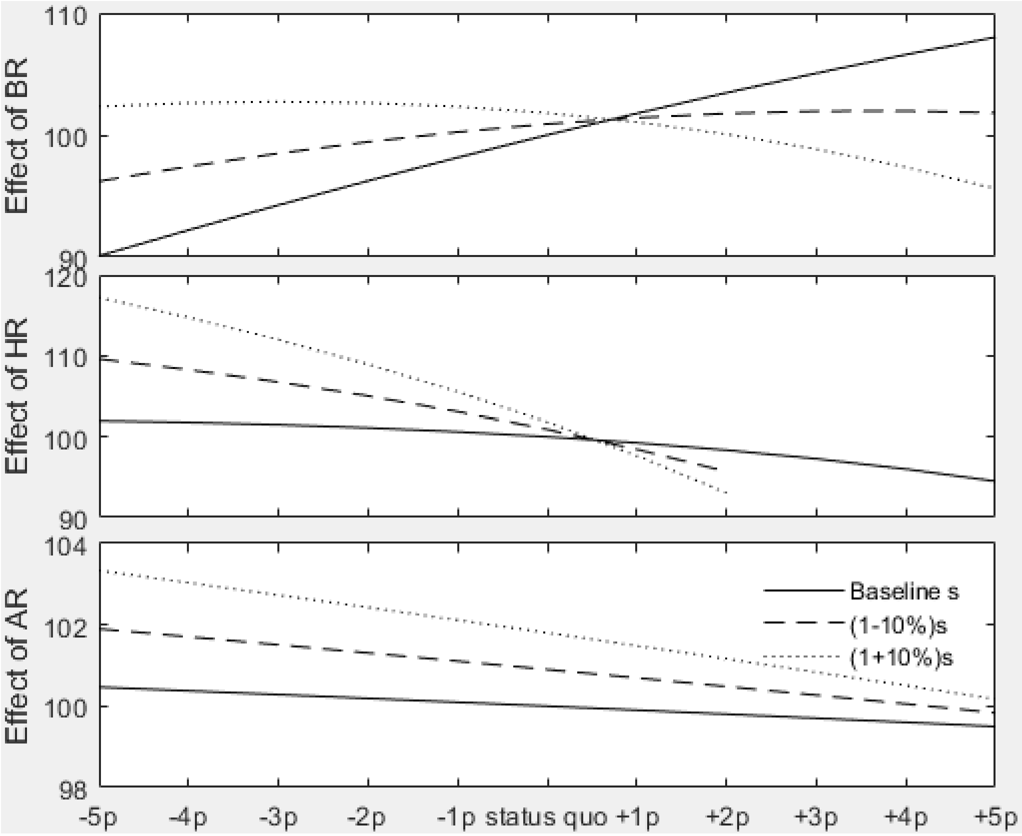

Figure D1: Robustness Check when the Elasticity of Substitution (s) Varies

As a robustness test, we have simulated the model with 10% deviations in either direction from the best estimate of the elasticity of substitution of the utility function. In the attached for the Appendix D, although the curvature in response to different rates of income tax differs, it can be seen that the qualitative conclusions are unchanged. With the two boundary elasticities, it is not possible to get an effect for HR beyond an increase of 2p.

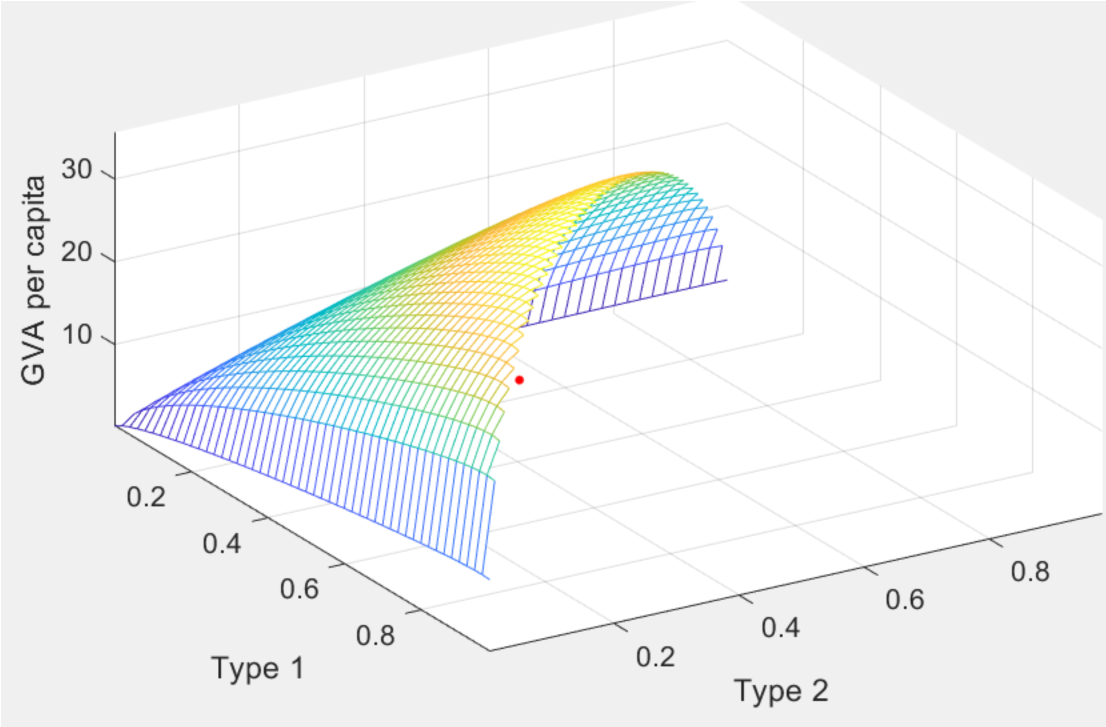

Figure D2: Visualisation of the Production Function (the red dot: status quo)

The figure represents the relation between taxpayer type shares and productivity. The horizontal axes are proportions of type 1 and type 2 and the vertical axis is the GVA per capita (in thousands). The red dot is the current position. If the shares of type 2 and type 3 are reduced by tax-induced migration, the average GVA will rise because of the lower productivity (wages) of type 1 taxpayers.