Optimal Federal Transfers Under Decentralised Regional Production Decisions

Jim Hancock*
Australian National University

February 28, 2015

Abstract

In recent years a number of authors have considered the optimal design of intergovernmental transfers in a federal system where information asymmetry limits the central government’s ability to implement a first-best transfer system. In most of those studies, regional populations are taken to be fixed. In this paper I consider the implications for policy design when there is information asymmetry and interregional population mobility. I consider three mobility scenarios: population immobility; complete, frictionless mobility; and mobility subject to friction in the form of “attachment to home”. I find that while it is optimal for a welfarist central government to allow an information rent to well-endowed regions when populations are immobile, the size of that information rent is diminished in scenarios with population mobility.

JEL classification: H23; H73; H77.
Keywords: fiscal federalism; intergovernmental transfers; migration; decentralisation; asymmetric information

This chapter presents a model of optimal resource allocation in a federal system, by which I mean a system comprising a central government and a number of regional governments with fixed jurisdictions. Each region has a population of homogenous worker-consumers who produce a composite commodity. Regions differ in an exogenous productivity parameter. Regional output can be consumed or used to purchase productivity increments.

I consider the implications of alternative population mobility scenarios, the implications of alternative information scenarios, and the interaction between these scenarios. The first mobility scenario is the case of fixed populations. The second scenario is the polar case where worker-consumers

*Draft. Not for quotation. I would like to thank Richard Cornes, Maria Racionero and Ronald Stauber, and participants in the ANU Applied Microeconomics Seminar for helpful comments. Comments gratefully received at jim.hancock@anu.edu.au.
are perfectly mobile and migrate in response to interregional differences in living standards. The third scenario is intermediate, and covers the case where regional populations are mobile but have a degree of attachment to home. In each of these scenarios I consider first the case where the centre has complete information regarding regional productivity and contrast it with the case where that information is private to regions and subject to strategic revelation.

There is a long literature dealing with optimal resource allocation in federal systems. Oates (1972) canvasses key issues that need to be resolved: the optimal size of regional jurisdictions, the optimal allocation of functional responsibilities and powers to particular types of jurisdictions, the optimal design of grants from the centre to regions, etc. In the literature, these issues are approached selectively, as an omnibus model quickly would become intractable. It is necessary to confine attention to a subset of the relevant factors.

Mobility has been seen as a potentially important phenomenon in federal systems at least since Tiebout (1956). If different local jurisdictions offer different levels of public services, then the residence choices available to a federal citizen effectively create a menu of public service mixes, and the citizen as a consumer then chooses from that menu by choosing which community to live in. In this view the citizen’s mobility implies a set of feasible location choices and thus the menu of public services available to him. Indeed Buchanan and Goetz (1972) argue that the Tiebout model treats location choices purely as a mechanism for choosing public service mixes from a menu, without allowing them any other significance, and in so doing excludes any geographic considerations that might be of interest. Of course if all citizens have the same preferences there is nothing to gain from differentiated service offerings, so the Tiebout model makes sense only in a setting where consumer preferences are heterogeneous. In contrast, in this chapter I confine attention to consumers who are homogenous or nearly homogenous. The only heterogeneity that will be considered is citizens’ ‘attachments to home’, which may be of varying intensity, and which are more of the character of frictions than differences in preferences. In a world with homogeneous worker-consumers, location choices will be affected by factors such as economies of scale in the provision of services, especially public services, and by congestion in consumption and production. The model in this chapter incorporates the effects of congestion in production activity: marginal products of labour are diminishing and this has implications for the distribution of workers and thus population.

Many papers in the fiscal federalism literature have taken populations as fixed. For instance, this was the approach of Oates (1972, pp. 95–104) in his pioneering discussion of optimal grant structures. This is primarily a matter of convenience, as it is generally recognised that most federal systems have substantial freedom of internal migration. Direct controls on
people’s location choices are rare. Somewhat more common are de facto controls, such as restrictions on employment and discriminatory access to housing, public services and other social programs. The extent of direct and de facto restrictions on internal migration is arguably very low in most federal countries, such as the United States, Canada, Australia and Germany. Indeed, each of the countries mentioned has constitutional protection for free internal migration. And even in those federal systems where there are some restrictions on internal migration, such as the European Union and China, insiders have much greater freedom to move than have outsiders.  

Thus the literature that assumes fixed populations is perhaps more pertinent to the analysis of transfers between nations with migration barriers, and even then there is a potential for migration decisions to respond to the economic settings adopted. Given the prevalence of federal systems that are characterised by substantial freedom of internal movement, and given the potential for movements to be affected by economic policies, we need to know whether the assumption of fixed populations can be justified as a convenient modelling simplification of little consequence or, instead, whether it has material consequences for the analysis. The answer is that assuming a fixed population does materially affect analysis, and we should therefore use models that incorporate mobility when we consider fiscal arrangements over polities with free or near-free internal migration.

In a static model, mobility is synonymous with location decisions. Within an economic model, it is natural to assume that people locate in response to the utility levels that can be realised in alternative locations. When the utility to an individual is higher in one location than another, it is natural for the individual to choose the high-utility location. When the congestion effects of additional population outweigh any scale economies, then the utility level in a location will be decreasing in its population.  

In this case, the equilibrium under free migration will be reached when the marginal private utilities are equalised across locations.

1 Citizens of EU member states have a right to reside in other member states long-term, subject to being economically active or being a student, and having sufficient resources not to burden the social services of their host country during their stay, or accompanying a person meeting these requirements (European Union 2009). In China, there are few direct restrictions on people relocating. But de facto controls come into play via the hukou system, which associates citizens with particular geographic areas, and then allows discriminatory access to housing and other social entitlements according to origin (see Wang 2008).

2 This leaves open the possibility that congestion effects might dominate scale advantages at some population levels but be subordinate at others.

3 Strictly speaking this needs to be qualified in recognition that some regions may be left vacant while having marginal private utility of occupancy below the level prevailing in occupied regions. In practice, in a federal system, the regions that we consider are sub-national jurisdictions, and we would rarely if ever see an empty jurisdiction. But within any country it certainly is possible to draw geographic boundaries that contain no residents. We can infer that the marginal utilities of living in such regions are too low to
Homogeneous worker-consumers’ location choices will be affected both by interregional differences in the cost of consumption items and in earnings from work. Economics of scale tend to decrease the tax price of public goods and therefore to attract individuals to regions with large populations, while congestion has a deterrent effect. And taking into account production, because consumers generally will live within the region in which they work, patterns of labour demand have a major bearing on the allocation of consumers across regions. Under competitive conditions, with diminishing marginal products of labour, regional labour demand must be such as to equate the marginal product of labour with the regional wage rate. Utilities are then equalised across regions when differences in marginal products of labour are just sufficient to offset differences in the minimum expenditures required to achieve the uniform utility level. If scale effects dominate in consumption opportunities, we would thus expect to see lower wages in regions with large populations, and vice versa.

The idea that inter-regional differences in living standards may induce population movements is not new in the federalism literature. For example, it is considered at least as far back as Buchanan (1950). It is the central focus of a number of articles in the early 1970s. Buchanan and Wagner (1970) incorporate it into a graphical exposition of migration decisions in a federal setting. In an analysis of efficient resource allocation in a federal system Flatters, Henderson, and Mieszkowski (1974) set out an optimisation problem in which workers’ utility levels are constrained to be the same across regions. A few years later, Broadway (1982) criticises work which neglects the influence of utility differentials on location decisions. He shows that under some circumstances mobility can remedy distortions to regions’ incentives that otherwise would lead to Pareto sub-optimal regional policy settings.

The assumption of frictionless mobility is itself strong, and can be seen as a polar opposite to fixed populations. A scenario in which populations are mobile, but with migrations being costly, seems more realistic. Mansoorian and Myers (1993) present such a model. They consider a case where residents have an “attachment to home” in their initial location, and are induced to move to another region only if there is a wage premium large enough to compensate for the disamenity of moving. In their model migrations may occur, but in general the equilibrium does not involve a complete equalisation of marginal productivities and utilities across regions.

The explicit use of a production function sets this model apart from most but not all studies of information asymmetry in federal systems. The most common model structure has regions choosing between private and

\[4\] Where residential and employment opportunities are 'close' to regional boundaries the potential to live and work in different jurisdictions is increased. Moreover, the preponderance of opportunities close to regional boundaries generally will be greater the smaller we draw the regions.
public goods according to preferences, which involves choices between consumption bundles on the production possibility frontier but not about the frontier itself. Some studies have own-source income varying across regions but still exogenous. For example, Laffont (1995) introduces variations in regional average income levels by allowing for varying proportions of high- and low-productivity firms, and Cremer and Pestieau (1997) have varying proportions of high- and low-income individuals. Cremer, Marchand, and Pestieau (1996) have exogenous incomes that are common within regions but vary across regions. In a model with fixed populations, Lockwood (1999) considers interregional productivity differences in the form of differences in an exogenous income level. In an extension he considers distortions to labour supply but still without any interregional mobility. In a model with fixed populations Bordignon, Manasse, and Tabellini (2001) consider a case where available time per worker is stochastic and exogenous for regions. Workers choose how much labour to supply and regional governments choose public good expenditures with associated tax requirements being met with a distorting labour income tax. My study is closest in spirit to Raff and Wilson (1997) who use a model with a neoclassical production function, mobile labour and immobile land, and a regional government that makes costly efforts to boost productivity.

The case where homogeneous, mobile individuals have no preferences for any particular region and have no costs of moving has very strong implications. Under these circumstances we can expect equilibrium in which the utility levels in different regions are equalised. This has an important feedback to the behaviour of regional governments. So long as regional governments each have the objective to maximise the welfare of mobile individuals domiciled within their own jurisdiction, it means that regional governments who otherwise might be seen as competing with one another to maximise regional living standards are actually joined in a common endeavour. No policy measure introduced by a region can raise its own living standards without raising living standards in all other regions. And a region cannot introduce a policy that impairs the living standards of another region without damaging its own living standards.

In reality, political tensions are a fact of life between constituent regions in federations. One might therefore be wary of a model which carries the implication that regions will have a unified view as to optimal federal policy settings. With some plausible alternative mobility assumptions, a degree of conflict between regions is introduced. Firstly, regions may contain immobile factors, and the returns to these will not be subject to the equalising forces that act on returns to mobile factors. If regions have an interest in the returns to the fixed factors domiciled within their jurisdiction, but not in fixed factors domiciled elsewhere, then there is a clear potential for them to disagree over the policy settings that are to be implemented. For example, one can easily envisage a regional decision maker taking a particular interest
in the position of owners of land in its jurisdiction, and seeking therefore to
maximise land rents in its own jurisdiction to the detriment of owners of land
elsewhere.\(^5\) Secondly, relocation is costly, so the adjustments that occur via
mobility are not instantaneous. Consequently, when a regional government
is able to capture an enhancement to local living standards it may persist
for several years. Thirdly, mobile citizens may be heterogeneous and may
have attachments to particular locations. Where this is so, mobility may
place limits on interregional differences in living standards without going so
far as to equalise them fully. In that case a region that succeeds in boosting
regional living standards at the expense of another region may retain part
of that advantage permanently.\(^6\)

1 The model

The model encompasses a single central government and a number of regions.
The welfare maximising centre has the power to make transfers—positive or
negative—to regions and can condition those transfers on the information
available to it. The question we are interested in is: What is the optimal
transfer policy for the centre to adopt? We consider the possibility that the
best course for the centre is a policy of *laissez faire*, with each region simply
left to its own resources, but we will see that in fact this is not generally
optimal.

We consider a finite number \(I\) of regions. The regions are populated
by consumers with homogenous preferences which can be represented by a
utility function over a composite consumption commodity. Within a region,
consumers can be represented by a representative consumer. Each consumer
provides a unit of labour for production. Regions produce output, with the
production level depending on region-specific labour input, an “endowment”
and a productivity-enhancing effort.

We will use this model to consider alternative scenarios for information
availability and population mobility. There are differences between the de-
cision problems arising under these scenarios in terms of the choice variables
available to decision makers and in the constraints on choices. The general
approach is to consider first a central planner model in which the central
government is the only decision maker, and then to consider a decentralised

\(^5\)An explanation for this form of objective is not needed for the flow of the argument,
but is at hand. If citizens of the federation can own land in any jurisdiction, but dispro-
portionately hold land in their own regions of residence—a form of heterogeneity—then the
regional decision maker may be able to boost residents’ living standards by boosting land
rents.

\(^6\)Another explanation is that regional decisionmakers, or those who choose them, con-
sistently and predictably misunderstand their region’s interests. There may be some
substance to this, but I shall confine analysis to the case where economic agents’ beha-
viour is guided by rational self interest, and I will leave aside an exploration of irrational
behaviour.
model in which regions take some choices with the centre’s role confined to imposing an interregional transfer policy. Consumers are never decision units in this analysis, although the behaviour of a regional decision maker is essentially the same as could be expected if the representative consumer were to take decisions.

In central planner mode, all of the choice variables in the scenario under consideration are at the discretion of the centre. This means that the centre makes decisions about productivity-enhancing effort at the regional level, about the allocation of output for consumption across regions and, in scenarios where there is population mobility, it also makes decisions about the allocation of population. The central planner case provides a useful reference point and is used to establish first-best allocations, as is common in literature of this type.

In decentralised decision mode, regions make decisions about effort levels and output, while the centre makes decisions about interregional transfers. This setting is a two-stage game: in the first stage the centre declares its transfer policy and in the second stage each region chooses its optimal strategy. We identify the optimal decision rules by backward induction. First, we find best responses for regions, conditional on the transfer policy, and secondly we find the centre’s optimal transfer policy subject to those best responses.

We commence by considering the case where regional populations are fixed. First, we identify the centre’s optimal transfer policy under complete information. Then we introduce information asymmetry. The centre still observes each region’s population and output but each region’s productivity parameter and productivity-enhancing effort level are private to itself. Secondly, we will investigate the implications of introducing population mobility to the model.

In the remainder of this section we set out the model in general terms—in a form that is able to encompass fixed and mobile populations and with and without complete information.

2 Regional problem

Our model contains a number of regional governments. In this subsection we will describe the decision faced by a region. In brief, the regional government’s problem is to maximise the utility of a representative agent. Utility depends on consumption of a single composite commodity. The consumption level is constrained by regional income, which is the sum of net output and a transfer (positive or negative) received from the centre. The region produces a single output using labour, a fixed endowment, and productivity-enhancing effort. The region must choose its effort level, and by implication its gross output level, optimally.
The region seeks to maximise the utility of its representative agent, and we will see that this requires maximising per capita consumption. The region’s per capita consumption depends on its gross output, the resources that it devotes to effort, a centrally imposed transfer from (or to) external regions, and its population.

2.1 Technology

I will specify the production technology so that the marginal productivity of labour is diminishing. Fixed marginal products are a harmless and convenient simplification in models where the population distribution is exogenous—see for example Lockwood (1999). But diminishing marginal products are a useful device when we want to consider mobile, endogenous populations. The point is illustrated in Figure 1, which depicts a situation where the total population is to be allocated across two regions, H and L, according to the rule that no person will locate in one region when her marginal product is higher in the other region. In panel (a), marginal products of labour are diminishing, and the equilibrium population allocation is given where marginal problems are equalised, which requires \( N_H > 0 \) and \( N_L > 0 \).

In panel (b) marginal products are constant and always higher in region H, with the result that the entire population locates in H and nobody locates in L. Thus the fixed marginal products model is unable to describe a world in which, as an endogenous outcome, the population is spread across locations. I want the model to admit dispersed population as an outcome, and therefore I incorporate diminishing marginal products into the model.

The gross output \( Y \) of the region is given by

\[
Y = F(N, \theta, E)
\]

---

7Lockwood (1999) constructs a model with fixed populations and exogenous income. In his study, “productivity” alludes to this exogenous income element. He has random variation in exogenous income, which is taken to be unobservable under imperfect information, and the centre then infers the productivity level, in two variants, either from government revenue raising or from government spending. In contrast, in our model gross income is observable, but depends on productivity and costly effort, neither of which is observable. The productivity level must be inferred from the observed output.

8Diseconomies of population can also be introduced with a congestible consumption commodity. This is the approach in much of the locational efficiency literature, the prime motivation for it being that the inclusion of local public goods introduces fiscal externalities that distort location decisions, providing one rationale for intergovernmental equalisation payments. For instance, Buchanan and Goetz (1972) consider the case where incomes are exogenous and location decisions are driven by the tax price of local public goods. And among other prominent works in the literature on location efficiency under complete information, Buchanan and Wagner (1970), Flatters, Henderson, and Mieszkowski (1974), Boadway and Flatters (1982), Wildasin (1986) and Wellisch (2000) discuss models which incorporate both diminishing returns in production and the scale economies associated with a congestible public good.
where $F : \mathbb{R}_+^3 \to \mathbb{R}_+$ is a production function embodying the extant technology.

Output depends, firstly, on the regional labour input $N \in \mathbb{R}_+$ of homogeneous worker-consumers. I assume that each member of the population supplies one unit of labour to the production process. Thus a region’s population and its labour supply are equal. Moreover, there is no labour-leisure choice.

Secondly, output depends on an exogenous endowment $\theta \in \mathbb{R}_+$ which can be interpreted broadly as anything (other than labour and effort) that is fixed and specific to the region. It could be a land resource, a mineral endowment, or some environmental characteristic that relates to production, such as rainfall. It might also relate to “soft” assets such as institutions, e.g. stability of government, public health, general levels of security and safety, etc.\textsuperscript{9} The endowment considered here is one-dimensional, which is a non-trivial restriction: real-world economies will always have multi-dimensional endowments and there is no reason to expect that they can be aggregated into a composite endowment. But to allow a multi-dimensional type space introduces complexities, both presentational and more substantial, which are best avoided for the present.\textsuperscript{10}

\textsuperscript{9}I deliberately avoid the usage “capital” because it introduces connotations that are undesirable in this context. First, this is a purely static model, whereas models with capital commonly incorporate dynamic processes around capital formation. Secondly, the endowment is exogenous, whereas capital is usually thought of as a choice variable or at least endogenous.

\textsuperscript{10}In the context of decision making under perfect information, the drawback of modelling a multi-dimensional endowment is only presentational. It means that there is a multiplicity of variables to consider. But in a model of decision making under incomplete information, a multi-dimensional type space may bring with it many-to-one linkages from the type space to observable variables, in which case the design of incentive compatible mechanisms...
Thirdly, output depends on the input of a productivity-enhancing effort $E \in \mathbb{R}_+$. I assume that a unit of effort costs one unit of output, and thus can define net output as

$$Z = Y - E.$$ 

For convenience, I will assume that all inputs to the production function are essential, i.e. that if the level of any input is zero then output will also be zero. This assumption is made for convenience. It ensures that for any region an output of zero is feasible, which will make the presentation easier when we come to consider a model with incomplete information. However, I believe that the key conclusions of this paper could still be reached without this assumption, albeit at a cost in terms of complexity and lost clarity.

**Assumption 1** (Inputs essential). For all $N \in \mathbb{R}_+, \theta \in \mathbb{R}_+$ and $E \in \mathbb{R}_+$

$$F(0, \theta, E) = 0$$
$$F(N, 0, E) = 0$$
$$F(N, \theta, 0) = 0$$

I assume that the production function has standard neoclassical concavity properties over its three inputs.

**Assumption 2** (Concave production technology). In the interior of $\mathbb{R}_+^3$ the aggregate production function $F$ is twice continuously differentiable with respect to $N$, $\theta$ and $E$. It has the properties

$$F_N > 0$$
$$F_\theta > 0$$
$$F_E > 0$$

and its Hessian, the matrix of its second derivatives, is negative definite.

**Remark.** *The negative definiteness of the Hessian implies that $F_{NN} < 0$, $F_{\theta\theta} < 0$ and $F_{EE} < 0$. Note also that while this Assumption is confined to the interior of $\mathbb{R}_+^3$, Assumption 1 imposes restrictions on the derivatives of $F$ on the boundary, e.g. $F_N = F_\theta = F_{NN} = F_{\theta\theta} = 0$ whenever $E = 0$, etc.*

Under Assumption 2, $dY/dE > 0$, $d^2Y/dE^2 < 0$, $dZ/dE > -1$ and $d^2Z/dE^2 < 0$ for $E \in \mathbb{R}_+$. Thus output and net output are strictly concave functions of effort on the interior of the domain.

I also adopt an assumption akin to one of the Inada conditions, that as effort becomes larger, its marginal product becomes arbitrarily small. This assumption helps to ensure the existence of a finite optimal effort level.

is more complicated. When more than one type maps to an observable, the message associated with that observable is not necessarily uninformative but it may be ambiguous.

**11** If one input is essential and none of it is used, then increasing another input will not increase output, and thus the marginal product of that other input must be zero and so too the marginal change of that marginal product.
**Assumption 3** (Limiting behaviour of marginal returns to effort).

\[
\lim_{E \to \infty} F_E = 0
\]

I have deliberately avoided any constant returns to scale assumption. Under constant returns to scale, an increase in output of some specified proportion can be achieved by increasing labour, endowment and effort by the same proportion. For generality, I wish to leave open the possibility that an equiproportionate increase in output, labour and effort might require a more-than or less-than proportional change in endowment. However, constant returns to scale are not excluded.

When we come to problems of incomplete information, the analysis will be made easier if we have the property of “increasing differences”, in the sense of marginal returns to effort, at a given effort level, being higher for agents with a higher endowment. Therefore I include the following assumption.

**Assumption 4** (Increasing differences). The production function has the property that

\[ F_{\theta_E} > 0. \]

Figure 2 illustrates the relationship between output, effort and net output that is implied by Assumptions 1-4. The graphs are shown for a common population and two different endowments, \( \theta_H \) and \( \theta_L \), where \( \theta_H > \theta_L \). Inter alia, they illustrate two key implications of the increasing differences property. Firstly, panel (a) shows that, at any particular effort level, the graph of output as a function of effort (panel a) is steeper for \( \theta_H \) than for \( \theta_L \). The slope of this graph is in fact the marginal product of an extra unit of effort, and we see thus that Assumption 4 is satisfied. Secondly, panel (b) shows a graph of the effort required to produce a given output level. We see that, for any particular effort level, the graph is always flatter for \( \theta_H \) than for \( \theta_L \). The slope of this graph shows the marginal effort increment to produce an extra unit of output. The extra effort is greater for a region with low endowment than for a region with high endowment.

In fact it will be helpful to adopt output as a control variable in lieu of effort. The reason is that when we come to consider incomplete information we will maintain the assumption that output is observable while effort is not, and it is much simpler to talk about our problem with an observable control variable. To accommodate this, define the effort function implicitly as

\[
G(N, \theta, F(N, \theta, E)) = E. \tag{1}
\]

---

12 These graphs embody a CES production technology with an elasticity parameter of less than one, but there are other functional forms that would satisfy the assumptions made here.
Figure 2: Output, effort and net output
The function $G$ is the inverse function of $F$ on its effort argument. It tells us the effort required to produce a specified output with a specified labour supply and endowment. Because $F$ is strictly monotonic and continuous on $\mathbb{R}_{++}^3$, $G$ is well-defined. Note however that $G$ is not defined for $N = 0$ or $\theta = 0$ because, in those cases, by the essentiality of inputs, an output of zero can be achieved with any effort level, and thus there is no unique corresponding effort level (and in contrast $F$ is defined on those boundaries). Moreover, Assumptions 2 and 4 have the following implications for the derivatives of $G$:

**Proposition 1** (Derivatives of effort function). In the interior of $\mathbb{R}_{++}^3$ the effort function $G$ is twice continuously differentiable with respect to $N$, $\theta$ and $Y$. It has the properties

$$
G_N < 0 \\
G_{\theta} < 0 \\
G_Y > 0
$$

and its Hessian, the matrix of its second derivatives, is positive definite.

**Proof.** See Appendix A. \qed

**Remark.** The first derivatives say that if labour or endowment are increased then less effort is needed to produce a given output, and positive definiteness says that the effort savings shrink in magnitude. The first derivative on output says that with given labour and endowment the production of more output requires more effort, and positive definiteness says that the required enhancement to effort increases with output.

### 2.2 Regional budget constraint

I turn now to the regional budget constraint. The region receives a transfer $T$ from the centre, which may be positive or negative. In determining a region’s transfer, the centre can take into account the region’s population, endowment and output, and thus we define the transfer function $T : \mathbb{R}_{++}^3 \to \mathbb{R}_+$ and set $T = T(N, \theta, Y)$.

Using $C$ to denote aggregate consumption, we can write the regional budget constraint as

$$
C \leq Y - E + T.
$$

Note that this budget constraint implies that, except for transfers by government, regional consumption is limited to regional net output and there is no diversion of regional net output away from the region. Thus the region

\footnote{We could include these cases by defining $G$ as a *correspondence* but it would greatly complicate the analysis without any commensurate advantage.}
cannot own any of the output of another region, nor can there be any external ownership of the output of the region. This is not a trivial restriction: in real-world situations inter-regional property income flows are common.\footnote{A more general model would allow for external claims on output via external ownership of factors of production, e.g. non-resident land holdings. The model here treats the productive capacity of the region as a type of common-property resource, available to those who reside in it but unavailable to those who reside elsewhere. This might for instance describe a situation where natural resource royalties accrue to the residents of the region where they are generated. Jeffrey Petchey and James Petchey (in progress) consider a case where the share of regional income going to labour is determined according to a marginal product rule with the residual forming a rent. A share parameter then determines how much of this rent is retained within the region. The two polar cases are full retention and zero retention. My analysis can be seen as a case of full retention.}

In what follows, consumer preferences are non-satiable, which implies that for any optimal choice the budget constraint will bind. Therefore we can define the aggregate consumption function  \( C : \mathbb{R}^3_+ \to \mathbb{R}_+ \) which relates aggregate consumption to population, endowment and output

\[
C(N, \theta, Y) = Y - G(N, \theta, Y) + T(N, \theta, Y). \tag{2}
\]

I introduce two technical restrictions on the transfer function. The first relates to continuity and differentiability; it has the effect of giving the consumption function desirable continuity and differentiability properties, which enables the use of calculus. However, it stops short of imposing concavity/convexity on the transfer function, which is useful because later, when we seek to identify optimal transfer functions, we leave open a wider range of solutions.\footnote{One of the key aims of our analysis is to identify optimal transfer functions. The assumption here can therefore be understood as restricting our choice of optimal transfer function to transfer functions that are continuously differentiable. This restriction is non-trivial, as there may in some circumstances be a superior transfer function that is not \( C^1 \) everywhere. Auerbach and Hines Jr. (2002) discuss this issue in the context of optimal taxation. Just how restrictive such an assumption is depends on whether or not there is a finite number of regions or a continuum of regions, and then in the case where there is a continuum, whether or not there are mass points in the type space. The consequences of this assumption, and in particular whether it excludes superior transfer policies, will be considered further when we come to consider the centre's choice of an optimal transfer policy.}

The second restriction is a regularity condition, that there is a finite upper limit on the derivative of transfers with respect to output of the transfer to a region. This restriction seems uncontroversial. In practice it is hard to see how a transferor could ever offer to a region a transfer function that did not have this property; the transferor would need to have unlimited resources at its disposal.\footnote{If the region were atomistic the restriction might be harder to justify on practical grounds. There is not necessarily a limit on the transfers that an external transferor could make to an atomistic region.}

**Assumption 5** (Differentiability and continuity of transfer function). The transfer function \( T \) is twice continuously differentiable at every point of its
domain.

**Corollary 2** (Continuity and differentiability of consumption function). The consumption function is twice continuously differentiable on $\mathbb{R}^3_+$. 

**Proof.** Because the consumption function (2) is a sum of three components, $Y$, $G$ and $T$, it will suffice to show that each of those components is twice continuously differentiable. $Y$ is invariant to $N$ and $\theta$ and thus is twice continuously differentiable. $G$ is twice continuously differentiable by Proposition (1). And $T$ is twice continuously differentiable by Assumption (5).

In the following it will be useful to represent the transfer function in terms of a fixed element and a marginal rate schedule. Firstly, define

$$ T(N, \theta) = T(N, \theta, 0), \quad (3) $$

i.e. $T$ is the transfer received when a region produces no output. Secondly, define

$$ t(N, \theta, x) = \frac{dT(N, \theta, Y)}{dY}, \quad (4) $$

i.e. $t$ is the marginal change in transfer per unit of marginal change in regional income when regional the income level is $Y$. Then because, by Assumption 5, the transfer function is differentiable, we can write it as

$$ T(N, \theta, Y) = T(N, \theta) + \int_0^Y t(N, \theta, x) \, dx. \quad (5) $$

This decomposition makes explicit the distinction between the factor influencing the region’s transfer that is in its control—output—and the factors that are not—the parameters of the transfer function $T$ and the $t$’s. The decomposition is very similar to that which is seen in the optimal taxation literature—see for instance. There, we think of a tax schedule comprising a component that is payable at zero income (possibly negative) and a set of marginal tax rates, one for each possible income level.

For the optimisation problems that follow, it is desirable to introduce the following regularity condition.\(^{17}\)

**Assumption 6** (Regularity condition for transfer function). For given $N$ and $\theta$ the marginal transfer rates $t(N, \theta, Y)$ have a supremum $\bar{t}(N, \theta)$.

\(^{17}\)This condition is I believe innocuous. Its purpose as a regularity condition is to ensure the existence of a solution to the regional maximisation problem to follow. Marginal transfer rates could be unbounded on the upside only if they became larger and larger as output increased, which could imply that there is no maximising output for a region. But under any plausible circumstances it would be impossible for a transferor to finance such a scheme of transfers. Hence it is innocuous to rule it out. I note also that this condition would be superfluous if we imposed concavity on the transfer function; it can therefore be seen as a necessary substitute that makes admissible a broader set of transfer functions.
2.3 Regional objective

I assume that the consumers in the region have homogenous preferences, and that these preferences have a Bernoulli representation $U = U(c)$, where $c$ is the consumer’s consumption, with $U' > 0$ and $U'' < 0$. Because each consumer’s labour supply is fixed at one unit there is no need to include a labour supply or leisure variable in the utility function.

The centre’s objective is to maximise the utility of a representative agent. In the current setting, this objective can be seen as a reduced form that is consistent with a number of more fundamental social preference structures, such as a Paretian objective, utilitarianism, a welfare objective with equity weight, and maximin. With the representative agent distributive concerns do not arise within the region, and in fact the regional objective is equivalent to maximising per capita consumption. It is helpful therefore to rewrite the regional budget constraint (2) in per capita terms

$$c(N, \theta, Y) = \frac{Y - G(N, \theta, Y) + T(N, \theta, Y)}{N}. \quad (6)$$

We can illustrate the representative agent’s preferences with indifference curves in per capita output-per capita transfer space. These curves can also be interpreted as iso-welfare curves for the regional government. I will consider the case where population is immobile, which is easier than the case of mobile population. Let $z$ and $\tau$ respectively denote per capita net output, and the per capita transfer. From (6), we have $c = z - \tau$. Noting that the representative agent has constant utility so long as per capita consumption is constant, we see that each indifference curve is also a curve of constant per capita consumption. Therefore the indifference curve for some consumption level $c_i$, in per capita output-transfer space, will be the graph of

$$\tau(N, \theta, Y) = c_i - z(N, \theta, Y).$$

Because each indifference curve holds $c_i$ constant, this graph is an inverted per capita net output function with its vertical position determined by the consumption level $c_i$. With population and endowment held constant, its shape will be a reflection, along the horizontal axis, of a graph having the shape seen in panel (d) of Figure 2. Indifference curves are illustrated in Figure 3 for consumption levels 0, $c_1$ and $c_2$ (where $c_1$ and $c_2$ are arbitrary subject only to $c_1 > 0$ and $c_2 > c_1$). The curves illustrate that, for any given output level, a higher transfer boosts consumption and thus lifts the regional representative agent onto a higher indifference curve. And all of these indifference curves have their minimal transfer at that output level which maximises regional net output: if a region fails to maximise net output, then it needs a larger transfer to make up for the shortfall.
2.4 Regional problem: fixed populations

I turn now to the regional optimisation decision in the case where population is immobile.

It would perhaps seem natural to adopt the level of productivity enhancing effort as the regional control variable. However, because there is a one-for-one correspondence between aggregate effort and aggregate output, we can also employ output as the regional control variable. And output has the advantage that, in the asymmetric information problem to follow, regional output will be taken to be observable to the central government whereas regional effort will not. Therefore we will use gross output as the control.

The regional problem can be written as

$$\max_Y U(c)$$

subject to the per capita budget constraint (6) and a non-negativity constraint on output. Dispensing with $N$ and $\theta$ as arguments on the grounds that they are fixed herein, the Lagrangian for this problem then is

$$\mathcal{L}(c, Y, \lambda, \mu) = U(c) + \lambda \left( \frac{Y - G(Y) + T(Y)}{N} - e \right) + \mu Y.$$  \hspace{1cm} (7)

Appendix B shows that a maximum exists, shows that first-order conditions must hold at the maximum, and identifies those first-order conditions. Let $Y^*$ denote the maximiser (which for the moment may be non-unique).

If $Y^*$ is an interior solution it must satisfy the first-order condition

$$1 - G_Y(Y^*) + t(Y^*) = 0.$$  \hspace{1cm} (8)
This is a necessary condition which says that at an interior maximum the marginal net output gain from a small change in output must be offset by a marginal reduction in transfer of equal magnitude.

When \( Y^* \) is a maximiser on the interior, the second-order condition

\[
t_Y(Y^*) \leq G_{YY}(Y^*),
\]

is also necessary, although it may not always hold away from \( Y^* \). If the inequality does hold strictly everywhere, then the first-order condition is sufficient for a unique global maximum. But this is too strong an assumption to impose at present: it is equivalent to ruling out transfer policies which permit rapid increases in the transfer rate over some output ranges. Rather than introducing this restriction I will, in the decision problems that we consider later, consider whether its absence gives rise to problems.\(^{18}\)

Figure 4(a) depicts a case in which the maximising output is an interior solution. It has three output levels that satisfy the first-order condition, which illustrates that while the first-order condition is necessary it is not a sufficient condition for a global maximum.

There is also the possibility that the maximum is a corner solution in

\(^{18}\)One reason for my hesitancy to impose strict concavity is that it has shown to be dubious in studies of optimal income taxation. Diamond (1998) and Saez (2001) conclude that optimal marginal income tax rate schedules may indeed be U-shaped for plausible choices of parameters relating to the social welfare function, labour supply elasticities and income distribution. While the context of my problem is different, those findings are cautionary, and they make me reticent to restrict the set of transfer functions so strongly. I have of course already imposed a restriction by requiring that the marginal transfer rate be a continuous function—this is a consequence of Assumption 5 which imposes twice continuously differentiable.
which case $Y^* = 0$. This can happen only if

$$1 - G_Y(0) + t(0) \leq 0. \quad (9)$$

For this situation to arise it must be the case that, at zero output, an increase in output leads to a fall in the combined sum of net output and transfer. An example of a corner solution is presented in Figure 4(b). Generally the corner maximum does not imply anything about the second derivatives, the exception being that if $1 - G_Y(0) + t(0) = 0$ then it is necessary that $Y(0) \leq G_Y(0)$.

It is interesting to ask: How do changes in the parameters of the transfer function affect the region’s optimal output choice? In a nutshell, the answer is that small changes in the transfer schedule affect optimal output only in so much as they affect the marginal transfer rate at optimal output. The following theorem provides a more precise answer. To assist with the presentation of the argument, it first addresses the (easier) case of a linear transfer schedule, and then addresses the (more comprehensive) case of a non-linear transfer schedule. In fact, the linear schedule is a special case of the non-linear schedule: the case in which the marginal transfer rate is uniform across output levels.

**Theorem 3 (Impact of transfers on optimal output).** Suppose that a region has a unique utility-maximising output level $Y^*$. Then:

1. If the region is subject to a linear transfer schedule, i.e. a schedule with the form

$$T(N,\theta,Y) = T(N,\theta) + t(N,\theta)Y,$$

then:

(a) The region’s optimal output $Y^*$ is independent of the fixed element of the transfer, $T(N,\theta)$, i.e.

$$\frac{dY^*}{dT(N,\theta)} = 0. \quad (10)$$

(b) If the non-negativity constraint on regional output is strictly binding at the optimal regional output—in which case regional output must be zero—then optimal output will be unaffected by a small change in the marginal transfer rate $t(N,\theta)$. If the non-negativity constraint is just-binding or slack at the optimal regional output then a small change in the marginal transfer rate will increase the optimal output level, i.e.

$$\frac{dY^*}{dt(N,\theta)} \begin{cases} = 0 & \text{if constraint on } Y^* \text{ is strictly binding}, \\ > 0 & \text{if constraint on } Y^* \text{ is just-binding or slack}. \end{cases} \quad (11)$$
2. If the region is subject to a non-linear transfer schedule, i.e. a schedule with the form

\[ T(N, \theta, Y) = T(N, \theta) + \int_0^Y t(N, \theta, x) \, dx, \]

then

(a) The region’s optimal output \( Y^* \) is independent of the fixed element of the transfer, \( T(N, \theta) \), i.e.

\[
\frac{dY^*}{dT(N, \theta)} = 0.
\] (12)

(b) If the region’s optimal output is a corner solution, it will be unaffected by a small change in any of the marginal transfer rates \( t(N, \theta, Y) \) applying at different output levels. If the optimal output level \( Y^* \) is an interior solution, small changes in the marginal transfer rates at any output levels other than the optimal output level will not change the optimal output level, but a small increase in the marginal transfer rate at the optimal output level will increase the optimal output level, i.e.

\[
\frac{dY^*}{dt(N, \theta, Y)} \begin{cases} 
= 0 & \text{if } Y^* \text{ is a corner solution,} \\
= 0 & \text{if } Y^* \text{ is an interior solution and } Y \neq Y^*, \\
> 0 & \text{if } Y^* \text{ is an interior solution and } Y = Y^*.
\end{cases}
\]

(13)

**Proof.** See Appendix C.

**Remark.** The Theorem is confined to the case where \( Y^* \) is a unique maximiser. If there were some other maximiser \( Y^{**} \), then with a change in the marginal transfer rate at \( Y^{**} \) it would be possible to increase per capita consumption by choosing some output level in the neighbourhood of \( Y^{**} \). It would be much messier to state the Theorem in a situation where the regional government is indifferent between two or more distinct output levels.

Theorem 3 applies only when there is a unique optimal output level. While this apparently involves some loss of generality, the loss is in fact not material, because we can anticipate that when the centre pursues its objective it will be motivated to choose its transfer schedule so that there is a strict global maximiser for the regional problem. I will justify this contention later, when we consider the centre’s optimal transfer function.
2.5 Centre’s problem under complete information

Now consider a federation with a central government and a number of regions each with its own government. The centre has the power to impose interregional transfers, and it does so with the aim of maximising social welfare according to its own set of social preferences. The regions each maximise the utility of a representative agent, following the model discussed in the previous Section 2.4.

An important practical and methodological question is the number of regions and the number of types to be modelled. The two are not the same: we could for instance have a large number of regions but just two types in the type space, or just two regions and many types in a single-dimensional type space, and so on. And, when we consider large numbers of regions or types, we can consider the limiting cases of continua of regions or types.

The case where there is a continuum of regions is appealing from a methodological perspective because one can then resort to the law of large numbers to treat the realised distribution of regional types as fixed. To illustrate, firstly, if there are two types “rich” and “poor”, each with probability \( \frac{1}{2} \), then with a very large number of regions the centre can design its transfer policy knowing that one half of the regions will be rich and one half will be poor. In contrast, in a finite population, resort to the law of large numbers is less than satisfactory. For instance, with just two regions, the centre knows that it may have either two rich regions (with probability \( \frac{1}{4} \)), or one rich region and one poor region (with probability \( \frac{1}{2} \)), or two poor regions (with probability \( \frac{1}{4} \)). Thus there are three probability-weighted scenarios to consider. Secondly, with a continuum of regions, each region is vanishingly small in the federation and therefore has the characteristics of a price taker. In contrast, if there are few regions, an individual region may exert some power over the centre’s policy settings, in which case it may not take the centre’s policy as given. For the present I will abstract from these considerations by considering a large number of regions.

All of this matters because the centre’s treatment of a region will depend on the characteristics of other regions. This interdependency arises because we will introduce a budget balance requirement over the centre’s transfer policy. That budget balance requirement will mean that the transfer to a region will depend not just on its own type, but also on the realised type profile of the other regions. If we have a continuum of types, then by the law of large numbers there is only one possible aggregated type profile for the other regions. Thus the continuum case is analytically convenient. But with a finite number of regions, and without recourse to the law of large numbers, we will need to specify the transfer to a region conditional on the realised characteristics of other regions, which is messy under private information because a number of possibilities will need to be considered. It also means for instance that regional decisions will be made under uncertainty, and we
then need to introduce assumptions about how the region decides under uncertainty. For example, is the region an expected utility maximiser? If we consider a problem in which that uncertainty does not arise, we can avoid addressing that question.\footnote{In many principal-agent problems with many agents, each relationship can be treated as independent of the others, e.g. if the principal is engaged in extracting for itself the maximum rent from each agent. But in this study, the budget constraint introduces a dependency between each centre-region relationship. The centre can increase its transfers to one region only by reducing transfers to the other regions collectively. For more on this issue see Chapter 7 in Fudenberg and Tirole (1991).}

These considerations nudge us towards an analysis employing a continuum of regions. But in many federal systems there is only a small number of regions. For example, considering the second tier of government, Australia has eight states or territories, Canada has 13 provinces or territories, Germany has 16 lander, Switzerland has 26 cantons, China has 34 administrative regions and the United States has 50 states. The United Kingdom has only three constituent regions with second-tier assemblies; possibly one or more new ones will be formed for England. The European Union has 28 member states. And even in those federations with a large number of regions, some regions can have a disproportionate share of the economy or population: California, for instance, accounts for about one-eighth of the US population, and New South Wales has about one-third of the Australian population. In studies where there are potentially millions of independent decision units involved each with immaterial influence over aggregate outcomes—e.g. optimal taxation—the continuum approach seems to be on solid ground. In the study of a federal system, we cannot have the same confidence that regions will have immaterial influence over aggregate outcomes. Therefore it is probably useful to have a model set up in which the decision units, the regions, are not vanishingly small in the national economy and polity, and a model with a finite number of regions serves this purpose.

Fortunately Cornes and Silva (2000, 2002) show a way to work with a finite number of regions while stepping around problems of uncertainty. If we assume that the type is known for every region but one, we can then consider the implications of private information for this one region without the complication of uncertainty regarding the circumstances of the other regions. In this case, the region with private information knows everything that there is to know, while the centre’s information is imperfect only in that it does not know one region’s type. This scenario is of course rather unrealistic, and it should be understood as an illustrative device rather than a model with general applicability.

We will commence by considering a finite number of regions. Suppose that there are $I$ regions indexed $i = 1, \ldots, I$. Let $N_i$, $\theta_i$, $E_i$, $Y_i$, $Z_i$ and $c_i$ denote, respectively, region $i$’s population/labour supply, endowment, effort, gross output, net output and per capita consumption.
I consider the case where the centre’s social preferences can be represented with a generalised utilitarian social welfare function. For ease of exposition we will work with the (discrete) utilitarian form

\[ W = \sum_{n=1}^{N} U(c_n) \]
\[ = \sum_{i=1}^{I} N_i U(c_i), \]  

(14)

where \( c_n \) is the per capita consumption of the \( n^{th} \) resident of the federation and where the second line makes use of the fact that there is a population of identical representative agents within each region.\(^{20}\) \(^{21}\)

I now consider two problems. The first is a central planner problem, in which the centre chooses all control variables—in this case the regional output and per capita consumption levels. This central planner problem is a conventional analytical device that is used to identify a first-best outcome that is useful as a reference point. The second problem is one in which regions choose their own output levels and the centre chooses a transfer policy. The transfer policy impacts on a region by adjusting its fiscal capacity \textit{and} by modifying the incentives faced by it. We will be interested to know whether the decentralised arrangement can deliver the first-best outcome reached in the central planner problem.

In each of these two problems I will constrain per capita consumption levels to be non-negative in every region. In the decentralised problem, this can be interpreted as a participation constraint.\(^{22}\)

\(^{20}\)Using the pure utilitarian form does not materially diminish the generality of the analysis so long as we consider choices made under certainty. The generalised utilitarian form is \( W = \sum_{n=1}^{N} \phi(U(c_n)) \), with \( \phi \) strictly increasing and concave. Where \( \phi \) has these properties, the solutions to the regional problems \( \max U(c) \) and \( \max \phi(U(c)) \) are identical, i.e. the cardinality of \( U \) matters for the centre’s decision but not for the regional decision, and therefore the solution to the regional problem is the same regardless whether we write it \( \max U \) or \( \max \phi(U) \). In the interests of notational simplicity I employ the pure utilitarian form. However, the cardinality of \( U \) might matter to regions in a setting with uncertainty, e.g. if they were expected-utility maximisers. When the cardinality of \( U \) does matter, as in a simulation, one common form of the s.w.f. is

\[ \phi(U) = \frac{U^\rho}{\rho}. \]

See Salanié (2011, p. 84).

\(^{21}\)We could express this in frequentist form but will leave aside that approach until we come to discuss a continuum of regions.

\(^{22}\)In a real-world application, regions’ outside options would need to be considered. For instance, the outside options of a state that wanted to exit the United States might be quite different from those of a member wishing to exit the European Union. Cornes and Silva (2000) analyse the design of federal transfers in a scenario in which participation constraints potentially do bind the centre’s choice.
Central planner problem

In central planner mode, the centre chooses the regional consumption and output levels to maximise welfare subject to the aggregate resource constraint, that aggregate consumption is equal to net output, and non-negative regional consumption and output. Thus the problem is

$$\max_{c_i, Y_i} \sum_{i=1}^{I} N_i U(c_i)$$

subject to

$$\sum_{i=1}^{I} N_i c_i = \sum_{i=1}^{I} [Y_i - G(N_i, \theta_i, Y_i)],$$

with $c_i \geq 0$ and $Y_i \geq 0$ for all $i$. The Lagrangian for the central planner’s problem then is

$$L = \sum_{i=1}^{I} N_i U(c_i) - \lambda \left[ \sum_{i=1}^{I} [Y_i - G(N_i, \theta_i, Y_i)] - N_i c_i \right] + \sum_{i=1}^{I} \mu_i Y_i + \sum_{i=1}^{I} \phi_i c_i,$$

and the first-order conditions for its choices of regional output and consumption levels are

$$Y_i : \quad \lambda[1 - G_Y(N_i, \theta_i, Y_i)] + \mu_i = 0, \quad i = 1, \ldots, I,$$

$$c_i : \quad U'(c_i) - \lambda = 0, \quad i = 1, \ldots, I$$

and

$$KT : \quad Y_i \geq 0, \quad \mu_i \geq 0, \quad \mu_i Y_i = 0.$$

Appendix D shows that a solution to this problem exists and that (with the exception of the case where aggregate net output is zero) the constraint qualification holds everywhere, which allows us to conclude that these first-order conditions must hold at the maximum. It also demonstrates the concavity of the Lagrangian, which establishes that there is a unique global maximum.

Let $Y_i^p$ denote the central planner’s optimal output choice for region $i$. For each region corner and interior solutions are a possibility. By the convexity of $G$ in output,

$$\text{if } 1 - G_Y(N_i, \theta_i, 0) \leq 0, \text{ then } Y_i^p = 0 \quad \text{and} \quad (15)$$

$$\text{if } 1 - G_Y(N_i, \theta_i, 0) > 0, \text{ then } 1 - G_Y(N_i, \theta_i, Y_i^p) = 0. \quad (16)$$

When the central planner chooses output levels according to these conditions, net output will be maximised within each region and thus in aggregate.
The first-order conditions for consumption tell us that per capita consumption levels must be equalised across regions, which is a standard first-best result in a utilitarian setting. Let $Z_i^P$ denote region $i$’s net output when it has output level $Y_i^P$, and let $c_i^P$ be the optimal per capita consumption level. Then equalisation of per capita consumption across regions requires $c_i^P = c^P$ for $i = 1, \ldots, I$ where

$$c^P = \frac{\sum_{j=1}^I Z_j^P}{\sum_{k=1}^I N_k}.$$  \tag{17}

Decentralised decision making

Consider now the case where decisions about output are decentralised to the regions, while the centre determines transfers—positive or negative—to the regions. The centre can condition its transfers on the information it has regarding each region’s circumstances, which in the current full information scenario is a region’s population, endowment and output.

We can conceive of this setting as a two-stage game: in the first stage the centre declares a transfer policy and in the second stage regions make choices about their output. We identify the optimal decision rules by backward induction. First, we describe best responses for regions, conditional on the transfer policy, and secondly we find the centre’s optimal transfer policy subject to those best responses.

As we saw in Section 2.4, a region’s objective, in reduced form, is to maximise per capita consumption, and thus its best response is to maximise per capita consumption given the centre’s transfer policy. The control variable that it uses for this purpose is its output; its population and endowment are fixed in the current scenario. Recall that $Y^*$ is the region’s best response output level, which depends on population, endowment and the transfer function parameters, i.e. $Y^* = Y^*(N, \theta, \mathcal{T})$. Also, let $c^* = c^*(N, \theta, \mathcal{T})$, $G^* = G^*(N, \theta, \mathcal{T})$ and $\mathcal{T}^* = \mathcal{T}^*(N, \theta, \mathcal{T})$ denote the corresponding per capita consumption, effort and transfers when a region with population $N$ and endowment $\theta$ produces its best response output.

Using (8) and (9) we see that a necessary condition for the regional best response under transfer policy $\mathcal{T}$ is

$$1 - G_Y(Y^*(\mathcal{T})) + t(Y^*(\mathcal{T})) \begin{cases} \leq 0, & \text{if } Y^*(\mathcal{T}) = 0, \\ = 0, & \text{if } Y^*(\mathcal{T}) > 0. \end{cases}$$

(The $N$’s and $\theta$'s are suppressed as arguments here, having in mind that they are fixed, to improve readability.)

Now we come back to the first stage of the game, the centre’s decision. The centre is to choose a transfer policy $\mathcal{T}$ which specifies the fixed and
marginal transfer parameters of the transfer function (5), i.e.

$$\mathcal{T} = \{T(N, \theta), t(N, \theta, Y) : N \in \mathbb{R}_{++}, \theta \in \mathbb{R}_{++}, Y \in \mathbb{R}_{+}\}. \quad (18)$$

Because the centre is able to observe each region’s population $N$, endowment $\theta$, and output $Y$, it is able to determine its transfers conditional on these variables.

Intuition leads us to expect that a policy of non-uniform lump-sum transfers will be optimal. We expect non-uniformity because differences in regions’ populations and endowments give rise to differences in productive capacity, and differentiated transfers will then be needed to reduce or eliminate regional disparities in per capita consumption levels. And lump-sum transfers are appealing in that they avoid distortions to regions’ output choices. For the transfer policy to have a lump sum character, each region’s transfer would need to be independent of its choices, meaning that transfers would need to be invariant to output. That much is intuition; we would like to verify it by demonstrating that, in a quite general setting where the centre chooses from a broad set of non-linear transfer schedules, it will choose a policy with lump-sum non-uniform transfers.

I will start by considering the centre’s problem when it chooses from a restricted set of transfer policies: transfer policies that are linear in output, which have the form

$$T(N, \theta, Y) = T(N) + t(N)Y. \quad (19)$$

In this case the centre must choose $T(N)$ and $t(N)$.

The centre seeks to maximise welfare, which depends both on the aggregate net output and on the distribution of that output for consumption. Aggregate net output depends on decentralised regional output decisions, and Theorem 3 tells us that a region’s best-response output is independent of the lump-sum component of its transfer. However, that Theorem also showed that with a linear transfer function the best-response effort level is dependent on the marginal response of transfers to effort, and therefore the centre’s choice of $t(N, \theta)$ is likely to have an impact on net output.

With a finite number of regions, the centre’s transfer policy strictly only needs to map those values of $N$ and $\theta$ that actually arise. If we define $T_i = T(N_i, \theta_i)$ and $t_i = t(N_i, \theta_i)$, then $i$’s transfer is

$$T_i = T_i + t_iY_i, \quad (20)$$

and the centre’s transfer policy is the parameters that it chooses for the transfer functions, i.e. $\mathcal{T} = \{(T_i, t_i) : i = 1, \ldots, I\}$.

We can write the centre’s problem as

$$\max_{\mathcal{T}} \sum_{i=1}^{I} N_i U(c_i^*(\mathcal{T}))$$

26
subject to the centre’s budget constraint

$$\sum_{i=1}^{I} (T_i + t_i Y_i) \leq 0,$$

and the participation constraints

$$c_i^* \geq 0, \quad i = 1, \ldots, I.$$

The Lagrangian for this problem is

$$\mathcal{L} = \sum_{i=1}^{I} N_i U(c_i^*) - \mu \sum_{i=1}^{I} (T_i + t_i Y_i^*) + \sum_{i=1}^{I} \phi_i c_i^*$$  \hspace{1cm} (21)

and the centre must choose the parameters of the transfer function, the $T_i$’s and the $t_i$’s, to maximise it.

Appendix ?? shows that the Kuhn-Tucker conditions must hold for this model (except for the uninteresting case where no region is able to produce a positive net output) and derives the first-order conditions for its solution. They are

$$t_i : \quad t_i = 0, \quad i = 1, \ldots, I; \hspace{1cm} (22)$$

which says that under the optimal transfer policy a region’s choice of output should have no impact on its transfer, and

$$T_i : \quad U'(c_i^*) - \lambda = 0, \quad i = 1, \ldots, I; \hspace{1cm} (23)$$

which says that the optimal transfer policy should equalise per capita consumption levels across regions. In fact, if the centre adopts this transfer policy, it will achieve the first-best optimum of the central planner, which gives rise to the following Theorem, variants of which are widespread in the fiscal federalism literature.\footnote{The earliest occurrence of such a result that I am aware of is Oates (1972, pp. 95-104). He shows that there is a set of matching transfers that induce a Pareto optimal outcome in a model of decentralised public service choices subject to externalities.}

**Theorem 4** (Decentralisation under complete information). Under complete information there is a decentralisation strategy, in which the centre chooses transfers to regions and regions choose their own output levels, that yields the first best outcome under central planning. That strategy is to implement the linear transfer policy described by the conditions in (22) and (23).

**Proof.** The condition in (22) requires the centre to set $t_i = 0$ for $i = 1, \ldots, I$. The region then chooses its best response output level $Y_i^{D}$. There are two
possibilities regarding the character of this best response. Either \( Y_i^D = 0 \), in which case we must have
\[
1 - G_Y(N_i, \theta_i, 0) \leq 0,
\]
or \( Y_i^D > 0 \), in which case we must have
\[
1 - G_Y(N_i, \theta_i, Y_i^D) = 0,
\]
with the convexity of \( G \) implying that
\[
1 - G_Y(N_i, \theta_i, 0) > 0.
\]
We conclude then that \( Y_i^D \) is determined by
\[
\begin{align*}
\text{if } 1 - G_Y(N_i, \theta_i, 0) &\leq 0, \text{ then } Y_i^D = 0 & (24) \\
\text{if } 1 - G_Y(N_i, \theta_i, 0) &> 0, \text{ then } 1 - G_Y(N_i, \theta_i, Y_i^D) = 0 & (25)
\end{align*}
\]
But these conditions are the same as for the central planner problem—see (15) and (16). And by the strict convexity of \( G \), (16) and (25) each have a unique solution, and thus a common solution. We conclude that \( Y_i^D = Y_i^P \) for \( i = 1, \ldots, I \). Output levels under decentralisation are the same as with central planning.

Equation (23) tells us that the centre must choose the lump-sum elements of transfers so as to equalise per capita consumption levels across the regions. The common per capita consumption level \( c_D \) can be shown to be the same as in the central planner problem as follows:
\[
c_D = \frac{\sum_{j=1}^{I} Z_j^D}{\sum_{k=1}^{I} N_k} = \frac{\sum_{j=1}^{I} Z_j^P}{\sum_{k=1}^{I} N_k} = c^P.
\]
The second line follows from the result that \( Y_i^D = Y_i^P \) for all \( i \) and therefore \( Z_i^D = Z_i^P \). The third line follows from (17). Thus the per capita consumption level in each region is the same as under central planning, and the first-best optimum of the central planner is achieved.

The optimal lump-sum transfer elements \( T_i^D \) are given by
\[
T_i^D = \frac{L_i}{\sum_k L_k} \sum_{j \neq i} Z_j^D - \frac{\sum_{j \neq i} L_j}{\sum_k L_k} Z_i^D, \quad i, j = 1, \ldots, I.
\]
An interpretation of this is that the transfer to region \( i \) is equal to its own population share of all the other regions’ output, less those other regions’
Figure 5: Decentralised output choices under complete information

population share of region $i$’s output. The ensuing consumption levels exhaust net output.

Figure 5 illustrates the optimal transfers for the case where regions are of two types, with respective endowments $\theta_H$ (“rich”) and $\theta_L$ (“poor”), with $\theta_H > \theta_L$. The indifference curve $U^0_L$ is the highest attainable for a poor region in the absence of transfers, and to reach it the effort level $Y^*_L$ must be chosen. Indifference curve $U^0_H$ is the highest attainable for a rich region in the absence of transfers, and is reached at effort level $Y^*_H$. The centre then institutes a lump sum transfer $T^D_L$ to poor regions which means that a poor region now can attain the indifference curve $U^D_L$ so long as it maintains effort at $Y^*_L$. It also implements a lump-sum (negative) transfer to rich regions, which means that the best a rich region can do is to maintain effort level $Y^*_H$ and reach the indifference curve $U^D_H$. The optimal transfers have the property that $U^D_L = U^D_H$. Because utilities are equalised, at zero output the two regions would require the same transfer; in the Figure they both touch the $\tau$-axis at $c^D$.

2.6 Centre’s problem under incomplete information

We turn now to a situation with incomplete information: some of the information that the centre needs to implement the first-best transfer policy is

---

24The positive transfer to poor regions and negative transfers to rich regions are implied by the centre’s budget constraint. That constraint also restricts the relative magnitudes of the transfers according to the population shares of the two types. The diagram is drawn for the case where population shares are equal, although that is by no means necessary to the discussion here.
private to regions. Specifically, assume that the centre is unable to observe regional endowments or effort levels, but can observe regional outputs and populations. In contrast each region has complete information regarding its own circumstances.

Under this information structure the centre must rely on each region to reveal its type \((N, \theta)\), or specifically the endowment element of it, as the centre knows its population. A region could do this by announcing its endowment to the centre (“direct revelation”). Alternatively, the centre might infer the region’s endowment by observing its behaviour, for example its choice of gross output (“indirect revelation”). Here, we will discuss the problem in an indirect revelation setting where the centre observes the regions output and makes inferences from it.

In addition, we assume that regions are willing to make inaccurate disclosures to the centre when it is in their own self-interest to do so.\(^{25}\) In this study a region’s self-interest is to maximise the utility of its representative agent, which requires it to maximise its per capita consumption. With indirect revelation, inaccurate disclosure involves the region adopting an output level that is consistent with the behaviour of a region with a different type. We say that the region then mimics that other type of region.

While the centre can observe each region’s gross output, this is not necessarily enough to infer the region’s endowment. The problem is that any particular gross output level could be the result of low endowment and high effort, or the result of high endowment and low effort, and the centre is unable to distinguish the two. However, if the centre takes into account the optimising behaviour of a region, and if there is a one-to-one relationship between optimal output levels and endowments, then the centre can observe output and infer the region’s endowment from this. The model that has been developed herein has this one-to-one property by virtue of Assumption 4, which imposes increasing differences on the production function.

In the case of complete information we identified the centre’s mechanism design problem as to choose a welfare-maximising menu of pairs \((T, (N, \theta))\) to be offered to regions. Under incomplete information, the centre must specify its menu in terms of observables, and therefore it offers pairs \((T, (N, Y))\). The challenge facing the centre now is to choose the offering that maximises welfare, subject to the incomplete information.

The centre proceeds by setting out a menu that has an offer for each of the possible types \((N, \theta)\). A region then “announces” its type, either

---

\(^{25}\)If regions could be relied on to disclose honestly under all transfer regimes, then the incomplete information problem would be trivial: the centre would simply ask regions to inform it. Regions might, for example, be ethically constrained to disclose honestly. Or it might be possible for the centre to audit the information provided by a region, in which case a sufficiently large sanction on false disclosure could induce truthful revelation by the region. The use of audit mechanisms in models of federal transfers is considered by Laffont (1995) and Cremer, Marchand, and Pestieau (1996).
explicitly or implicitly. With explicit announcement, a region of size $L$ makes an announcement $\hat{\theta}$ and the centre then checks that it produces the output consistent with this, $Y^*(N,\hat{\theta})$. With implicit announcement the centre observes the regional output level and population, and then infers the announcement $\hat{\theta}$. For this implicit approach to be viable, it is necessary that there be a unique $\theta$ associated with any $Y^*(N,\hat{\theta})$. Moreover, with explicit announcement, observing actual output will be an unsatisfactory verification mechanism if some output level is optimal for two different endowment levels. These difficulties are avoided if we confine ourselves to problems in which there is a one-to-one relationship between endowment and optimal output. A sufficient condition for this is that the menu be designed in such a way that $Y^*(N,\theta)$ is a strictly monotonic function of $\theta$.

With a region’s endowment private to itself, under some transfer policies it will be first-best for that region to make an untruthful report. For example, if the centre were to set its transfer to a region according to the region’s own report of its endowment, and independent of the regional choice of effort, then it would be rational for the region to report whatever endowment would give it the largest transfer. And with all regions behaving in this way, the centre’s budget balance requirement would force it to set this largest transfer at zero, i.e. it would not be feasible to vary transfers to compensate for innate productivity differences. But with such a transfer policy, per capita consumption levels cannot be equalised, and the first-best optimum seen under complete information cannot be attained.

The revelation principle tells us that the centre can confine itself to offers that will induce truthful revelation. Offers that do not induce truthful revelation can be excluded on the basis that they cannot be implemented. To bring about truthful revelation, the centre needs to make offers to regions that are “incentive-compatible” in the sense of having the property that the best strategy available to a region is truthful revelation of its type. We can express this in the incentive compatibility constraint

$$U(c^*(\hat{\theta}_m|N,\theta_m)) \geq U(c^*(\hat{\theta}_n|N,\theta_m)) \text{ for all } \hat{\theta}_m, \hat{\theta}, \theta_m \in \Theta$$

where $^\dagger$ denotes a declared type and $\Theta$ is the type space. where $\hat{\theta}_n$ denotes that the type is declared to be $\theta_n$, $\theta_m$ denotes that the type actually is

---

$^{26}$For discussions of the revelation principle see Chapter 7 in Fudenberg and Tirole (1991), the appendix to Ch. 1 in Laffont and Tirole (1993), Ch. 14 in Mas-Colell, Whinston, and Green (1995), Ch.2 in Laffont and Martimort (2002), Ch. 2 in Salanié (2005) and Ch. 20 in Binmore (2007). Binmore observes that the revelation principle doesn’t provide a magical way of getting to a first best outcome …The revelation principle says only that, if something can be done at all, then it can somehow be done by asking people to reveal their true types …but the limitations on what a principal can know or do in a given situation usually force a second-best outcome on the designer of an optimal mechanism [pp. 574-575]
\(\theta_m\), \(c^*(\hat{\theta}_m|\theta_n)\) is the maximum per capita consumption available to a region of type-\(n\) when it declares itself to be a type-\(m\) region, and \(\Theta\) is the type space. When \(m = n\), revelation is truthful, and the purpose of the incentive compatibility constraints is to ensure that truthful revelation is optimal for every type.

At this point we will simplify the problem by restricting the type space to just two values: \(\Theta = \{\theta_H, \theta_L\}\). We will also confine attention to the case where all regions have the same population \(N\). With these restrictions there are just two incentive compatibility constraints

\[
U(c^*(\hat{\theta}_H|N,\theta_H)) \geq U(c^*(\hat{\theta}_L|N,\theta_H)) \tag{26a}
\]

and

\[
U(c^*(\hat{\theta}_L|N,\theta_L)) \geq U(c^*(\hat{\theta}_H|N,\theta_L)) \tag{26b}
\]

In the simple two-type setup considered here, there are two deviations from truth-telling to consider. The first case is that a low type pretends to be a high type. We will set out an optimisation problem on the assumption that mimicry of this form will not occur, which allows us to dispense with (26b) as a constraint in the program to follow; later we will confirm that the solution to that program is compliant with this assumption. The reason why it will not occur is that our solution will have the character that high types receive negative transfers whereas low types receive positive transfers. Under these circumstances, if a low type were to mimic a high type it would, firstly, receive a less generous transfer and, secondly, distort its effort level away from that which maximises net output. These two factors each diminish the region’s per capita consumption and thus their combined effect is to diminish per capita consumption. Therefore, for a low type to mimic a high type it would be inconsistent with its objective of maximising per capita consumption, and the associated incentive compatibility constraint is redundant. The second deviation from truth-telling is that a high type mimics a low type. In this case mimicry entails distorting the effort level, which is adverse for consumption, but it also leads to a more generous transfer which boosts consumption. If the increase in the transfer from mimicry is large enough, mimicry is beneficial to the high-endowment region. Therefore the incentive compatibility constraint for the high type is essentially that the difference in the transfers received by high and low types should not be so large as to induce false revelation by the high type.

Figure 6 shows the situation when the centre uses the first-best optimal output levels and transfers (taken from Figure 5) as the basis for its menu. As there are only two types in this example, that menu includes only two offers: \((Y^P_H, T^P_H)\), intended for a high endowment type, and \((Y^P_L, T^P_L)\), intended for a low endowment type. If the region has a high endowment and chooses the offer intended for it, it attains the indifference curve \(U^D_H\). But in fact it
can do better by choosing the offer intended for the low endowment type, because it then attains the higher utility level of indifference curve $U^M_H$. 27The high type incurs a cost when it mimics the low type, arising from foregone consumption when it distorts its output from the optimal level $Y^D_H$, but this cost is more than compensated by the extra transfer revenue that it receives. In conclusion, under a menu based on the first-best transfer policy we cannot expect a high type to declare itself truthfully. Moreover, this problem will arise whenever the offer for the low type lies in the interior of the hatched region of Figure 6. To remedy this problem, the centre needs to modify its offers from a first-best basis, and having in mind the budget balance constraint it must simultaneously improve the offer available to the high type and make less appealing the offer for the low type. There are many incentive feasible menus, and the challenge is to choose that which satisfies incentive compatibility at minimum sacrifice of the welfare level achieved under the first-best optimum.

For illustrative purposes, we will follow Cornes and Silva (2000, 2002). They begin their analysis by considering a minimalist form of informa-

\footnote{This observation is accurate for the indifference curves shown here. However, it is possible to have a situation where the first-best optimum is incentive compatible. This would be true when the difference in transfers offered to the two types is insufficient to compensate a high type for the costs of distorting its output choice to mimic a low endowment type. This case is of limited relevance, because when there is a continuum of types, a menu that offers the output-transfer combinations under the first-best cannot be incentive compatible. And in a real world setting, productivity levels are likely to be distributed across a continuous type space, meaning that the conclusions of the continuum case are likely to apply. The two type case is used not for its realism but because it is a useful illustrative device.}
tion asymmetry and thus draw out the consequences of private information without some of the clutter that is introduced when we consider a model with many types and many regions with private information.

I consider a case where there are \( I \) regions which are of only two types. Rich regions have endowment \( \theta_H \) while poor regions have endowment \( \theta_L \), where \( \theta_H > \theta_L \). The centre is unable to observe the type of any particular region but knows that the proportion of the population living in rich regions is \( \pi_H \) and that the proportion living in poor regions is \( \pi_L \); with just two types \( \pi_L = 1 - \pi_H \). Admittedly this knowledge structure is rather contrived, but it is a useful first step as its simplicity makes it easier to identify consequences of information asymmetry without the complications of, for instance, a stochastic setting.

We will now consider the case of a linear transfer policy. We can write the centre’s problem as

\[
\max_T \sum_{i=H,L} N \pi_i U(c_i^*(T))
\]

subject to the centre’s budget constraint

\[
I \sum_{i=H,L} \pi_i(T_i + t_i Y_i^*) \leq 0,
\]

the incentive compatibility constraint (26a)

\[
U(c^*(\theta_H|N, \theta_H)) \geq U(c^*(\theta_L|N, \theta_H))
\]

and the participation constraints

\[
c_H^*, c_L^* \geq 0.
\]

The Lagrangian for this problem is

\[
\mathcal{L} = \sum_{i=1}^{I} N_i U(c_i^*) - \mu \sum_{i=1}^{I} (T_i + t_i Y_i^*) + \sum_{i=1}^{I} \phi_i c_i^*
\]

and the centre must choose the parameters of the transfer function, the \( T_i \)’s and the \( t_i \)’s, to maximise it.

Appendix ?? shows that the Kuhn-Tucker conditions must hold for this model (except for the uninteresting case where no region is able to produce a positive net output) and derives the first-order conditions for its solution. [Incomplete.]

A Proof of Proposition 1

Firstly, introduce a more compact notation. For convenience, define \( \mathbf{a} = [N \ \theta \ E]^T \) and \( \mathcal{F}(\mathbf{a}) = [N \ \theta \ F(N, \theta, E)]^T \), and then rewrite (1) as \( G(\mathcal{F}(\mathbf{a})) = E \).
Secondly, if we differentiate \( G(F(a)) = E \) with respect to \( a \), we see that
\[
[D_a G(F(a))] = [0 \ 0 \ 1],
\]
and then using the chain rule on the left-hand side, we have
\[
[D G(F(a))][D F(a)] = [0 \ 0 \ 1].
\]
The expansions of the two derivatives on the left-hand side are
\[
[D G(F(a))] = [G_1 \ G_2 \ G_3] \quad \text{and} \quad [D F(a)] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ F_1 & F_2 & F_3 \end{bmatrix},
\]
and therefore we have \( G_1 = -F_1/F_3 < 0, \ G_2 = -F_2/F_3 < 0 \) and \( G_3 = 1/F_3 > 0 \), where the inequalities follow from \( F_1 > 0, F_2 > 0 \) and \( F_3 > 0 \) on the interior of \( \mathbb{R}^3_+ \) (Assumption 2). Thus the part of the Proposition relating to the first derivatives is demonstrated.

Thirdly, using the chain rule again we have
\[
[D_a [D G(F(a))][D F(a)] + [D G(F(a))][D^2 F(a)] = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},
\]
and thus
\[
[D^2 F(a)]^T[D G(F(a))][D F(a)] = -[D G(F(a))][D^2 F(a)] \quad (28)
\]
Expressing \( D^2 F(a) \) is problematic as \( [D F(a)] \) is already a 3 × 3 matrix. However, if we write
\[
[D G(F(a))][D F(a)] = G_1[D F(1)(a)] + G_2[D F(2)(a)] + G_3[D F(3)(a)]
\]
where \( [D F(i)(a)] \) is the \( i \)th row of \( [D F(a)] \), then we see that
\[
[D[D F(1)(a)]] = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},
\]
\[
[D[D F(2)(a)]] = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},
\]
and
\[
[D[D F(3)(a)]] = \begin{bmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{bmatrix} = [D^2 F(a)],
\]

35
so that
$$[DG(F(a))][D^2F(a)] = G_3[D^2F(a)].$$
If we substitute this into (28) we have
$$[DF(a)]^T[D^2G(F(a))][DF(a)] = -G_3[D^2F(a)].$$

Visual examination of $[DF(a)]$ shows that it is of full rank and invertible, since $F_3 > 0$ always. Therefore we can write
$$\begin{pmatrix} G_{11} & G_{12} & G_{13} \\ G_{21} & G_{22} & G_{23} \\ G_{31} & G_{32} & G_{33} \end{pmatrix} = [D^2G(F(a))] = -G_3[[DF(a)]^{-1}]^T[D^2F(a)]^T[D^2F(a)].$$
The $3 \times 3$ matrix $[[DF(a)]^{-1}]^T[D^2F(a)][DF(a)]^{-1}$ is symmetric and negative definite. This symmetry follows from the symmetry of $[D^2F(a)].$ And because $[DF(a)]^{-1}$ is of full rank, then for any $v \neq 0$ we must have $[DF(a)]^{-1}v \neq 0.$ But then by the negative definiteness of $[D^2F(a)]$ we must have $v^T[[DF(a)]^{-1}]^T[D^2F(a)][DF(a)]^{-1}v < 0$ for all $v \neq 0,$ and thus negative definiteness is established. Finally, because $G_3 > 0,$ we conclude that $[D^2G]$ is positive definite. Thus the part of the Proposition relating to second derivatives is demonstrated.

**B Regional optimisation problem**

This Appendix addresses the regional optimisation problem with fixed population. It demonstrates the existence of a solution to the regional optimisation problem, considers the constraint qualification to identify those parts of the domain for which the first-order conditions of the Kuhn-Tucker procedure are necessary for a maximum, and then identifies those first-order conditions and second-order conditions for the transfer function. I suppress $N$ and $\theta$ as arguments throughout because they are given for this problem.

**Existence of a maximum**

Here I demonstrate the existence of a utility maximising output level. For this purpose I draw on the Weierstrass Theorem, which says that a continuous function with a compact domain in the reals attains a maximum on that domain—see Sundaram (1996) pp. 90–91.

The objective function $U(c(Y))$ is continuous on $Y$ since $U$ is continuous and $c = C/N$ must also be continuous with respect to $Y$ since $N$ is a non-zero constant and $C(Y)$ is continuous by Corollary 2.

The domain is compact if it is closed and bounded. But regional output may take values on $\mathbb{R}_+$, which is closed but not bounded. However, we can
get around this problem using the second condition in Assumption 6, which tells us that the marginal transfer rates have a supremum \( \bar{t} \). Note that

\[
\frac{dU(c(Y))}{dY} = U'(c(Y)) \frac{1 - G_Y(Y) + t(Y)}{N}.
\]

First, consider the case where, at \( Y = 0 \), we have \( 1 - G_Y(0) + \bar{t} \leq 0 \). Because \( G \) is convex \( G_Y \) is strictly increasing, so in this case we must have \( 1 - G_Y(Y) + \bar{t} Y < 0 \) for all \( Y > 0 \). And since \( U''(c) > 0 \) for all \( Y \), it must be the case that \( dU/dY < 0 \) for all \( Y > 0 \). Therefore \( Y > 0 \) cannot be optimal for this region. We declare \( Y = 0 \), and can say that no \( Y > Y \) can be a maximiser.

Secondly, consider the case where, at \( Y = 0 \), we have \( 1 - G_Y(0) + \bar{t} > 0 \). Define \( Y \) as the solution to \( 1 - G_Y(Y) + \bar{t} = 0 \); by the strict convexity of \( G \) we know that it exists and is unique. For any \( Y > Y \), by the strictly increasing character of \( G_Y \) we must have \( 1 - G_Y(Y) + \bar{t} < 0 \). And since \( t < \bar{t} \) everywhere, we must also have \( 1 - G_Y(Y) + t < 0 \) and therefore \( dU/dY < 0 \) when \( Y > Y \). Therefore no \( Y > Y \) can be a maximiser.

All of this means that we can restrict the domain of \( U(c(Y)) \) from \( \mathbb{R}_+ \) to \([0, \bar{Y}]\) without changing the solution set for the maximisation problem. The solution to the problem on this restricted domain is therefore identical. But because the domain is now bounded, while still closed, it is now compact, and we can invoke the Wierstrass Theorem to conclude that a maximum exists.

**Constraint qualification**

Here we check whether—and where—the constraint qualification for the Kuhn-Tucker method holds. Define

\[
h_\lambda(c,Y) = \frac{Y - G_Y(Y) + \bar{t} Y}{N} - c.
\]

(i.e. the function which is used to express the regional budget constraint) and

\[
h_\mu(c,Y) = Y
\]

(i.e. the function that is used to constrain output to non-negativity). Let \( \mathbf{h_E} \) be a column vector of effective (i.e. binding) constraints. In the case where \( Y > 0 \), the only effective constraint is the regional budget constraint, and then we have

\[
[\mathbf{Dh_\lambda}] = [\mathbf{Dh_\mu}] = \begin{bmatrix} -1 & \frac{1}{N}(1 - G_Y(Y) + t(Y)) \end{bmatrix}.
\]

Here, \( \text{rank}([\mathbf{Dh_\lambda}]) = 1 \), which is the equal to the number of effective constraints, so the constraint qualification is met. If \( Y = 0 \), then we have two effective constraints and thus

\[
[\mathbf{Dh_\lambda}] = [\mathbf{Dh_\mu}] = \begin{bmatrix} -1 & \frac{1}{N}(1 - G_Y(Y) + t(Y)) \\ 0 & 1 \end{bmatrix}.
\]

37
Here rank(\(|\mathbf{Dh_\lambda}|\)) = 2, which again is equal to the number of effective constraints, so again the constraint qualification is met. Since the constraint qualification holds in both cases, the first-order conditions of the Kuhn-Tucker Theorem also must hold.

**Critical points**

Here we identify the first-order conditions for critical points of the Kuhn-Tucker problem. From the Lagrangian (7) they are

\[
c : \quad U'(c) - \lambda = 0
\]

\[
Y : \quad \frac{1}{N}(1 - G_Y + t) + \mu = 0
\]

\[
\lambda : \quad \frac{Y - G + T}{N} - c = 0
\]

**KT** : \(Y \geq 0, \quad \mu \geq 0, \quad \mu Y = 0\).

Consider separately the cases where \(\mu Y = 0\) is satisfied slack and where it is binding.

With non-negativity satisfied slack or just-binding, we have \(\mu = 0\), and then by the first-order condition for \(Y\) we must have

\[1 - G_Y + t = 0,
\]

i.e. a small change in output would lead to no change in aggregate consumption. If some \(Y^*\) satisfies this first-order condition, then to be a local maximum it must be the case that \(\frac{d^2 \mathcal{L}}{dY^2}|_{Y^*} \leq 0\), with the inequality strict for a unique local maximum. Since

\[
\frac{d^2 \mathcal{L}}{dY^2} = \frac{1}{N}(-G_{YY} + t_Y),
\]

then at a local maximum, by the strict positivity of \(\lambda\) and \(N\), it is necessary that

\[t_Y \leq G_{YY},
\]

with the inequality holding strictly for a strict local maximum. By Proposition 1, \(G_{22}\) is strictly positive, so what this condition means is that the marginal transfer rate cannot be increasing too fast at a utility-maximising output level (and there is nothing to rule out that it be decreasing).

With non-negativity strictly binding, we have \(\mu > 0\), so by the first-order condition for \(Y\)

\[1 - G_Y + t < 0
\]

must hold. For a corner solution, the marginal output gain when output is increased from zero must be less than the effort cost of that output gain net of the associated increment to the region’s transfer.
C Proof of Theorem 3

D Central planner output choice under complete information

Existence of a maximum

Firstly, we want to establish the existence of a maximum for this problem. For this purpose, we will use the Weierstrass Theorem, which says that a maximum exists if our objective function (the Lagrangian in this case) is continuous and the feasible set for the choice variables is compact.

The Lagrangian is a sum of functions that are continuous in the $Y_i$’s and $c_i$’s and is therefore itself continuous.

However, the choice set as specified is not compact: it is closed but it is not bounded as there is no upper limit imposed on the $Y_i$’s and $c_i$’s. However, I will use the other constraints in the problem to show that we can, without any restriction to the feasible set, impose upper bounds on the choice set—i.e. stipulate that there are values which the $Y_i$’s and $c_i$’s cannot exceed. We can then invoke Weierstrass to conclude that a maximum does exist on this restricted domain. The maximum must of course be in the feasible set, and because the feasible set is unchanged from our initial problem, we conclude that our initial problem (with unbounded domain) must also have a maximum. All that remains then is to show that we can “compactify” the domain without impinging on the feasible set, which I now do.

To identify an upper bound on output choices, note that the production technology implies that: at a regional output of zero, the region’s net output is zero; when we increase output, net output increases and eventually reaches a maximum; beyond that maximum, net output is decreasing in output; at some finite output level the region’s net output falls to zero; beyond that output level the region’s net output is negative. Let $\bar{Y}_i$ be the output level at which region $i$’s net output falls back to zero, and let $\bar{Y}_{\text{max}}$ be the maximum of the $\bar{Y}_i$’s. I assert that it is not feasible to produce an aggregate output of $I\bar{Y}_{\text{max}} + \epsilon$ for any $\epsilon > 0$. To do so would entail a negative aggregate net output. But if aggregate net output were negative, then by the aggregate budget constraint per capita consumption would be negative in at least one region, which is inadmissible. Because the feasibility constraints imply that no region can have output greater than $I\bar{Y}_{\text{max}}$, we can without impact on the feasible set, confine the choice of regional outputs to the set $\{\prod_{i=1}^I [0, I\bar{Y}_{\text{max}}]\}$, which is compact.

To identify an upper bound on per capita consumption choices, note that because the set of feasible outputs is compact, and the production technology is concave, there is a maximal aggregate net output level, call it $\sum Z^*$. Letting $S$ designate the region with smallest population, the max-
imum achievable per capita consumption level in $S$ would be achieved by producing the maximal net output and allocating it all to $S$, in which case $c_S = \sum Z^*/N_S$. And because no other region has a population smaller than $N_S$, there is no feasible allocation that could achieve a per capita consumption level higher than $c_S$ in any region. Therefore we can, without any impact on the feasible set, restrict the domain of the consumption variables to $\{\prod_{i=1}^I [0, c_S]\}$, which is compact.

**Constraint qualification**

Let $Y = (Y_1, \ldots, Y_I)'$ and $c = (c_1, \ldots, c_I)'$. Form the constraint vector

$$h(Y, c) = \begin{bmatrix} \sum_{i=1}^I [Y_i - G(N_i, \theta_i, Y_i) - N_i c_i] \\ Y_1 \\ \vdots \\ Y_I \\ c_1 \\ \vdots \\ c_I \end{bmatrix}$$

which has dimension $2I + 1 \times 1$. Now form the derivative matrix

$$[Dh(Y, c)] = \begin{bmatrix} 1 - G_Y(1) & \ldots & 1 - G_Y(I) & N_1 & \ldots & N_I \\ 1 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 1 & 0 & \ldots & 0 \\ 0 & \ldots & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & 0 & \ldots & 1 \end{bmatrix}$$

(29)

where $G_Y(i)$ denotes $G_Y(N_i, \theta_i, Y_i)$; the matrix has dimension $2I + 1 \times 2I$. Now form the derivative matrix for the effective constraints $[Deh(Y, c)]$ by deleting any non-binding constraints from $[Dh(Y, c)]$.

Consider first the degenerate case where all of the $2I + 1$ constraints are binding, which means all $Y_i = 0$ and $c_i = 0$ for all $i$. In that case the derivative matrix of the effective constraints is exactly as in (29). The last $2I$ rows of that matrix clearly are independent, and thus the matrix has rank of $2I$. But this is less than the number of effective constraints, $2I + 1$, and thus the constraint qualification is not met.

Now consider the case where the constraint for some $Y_i$ is not binding. To generate $[Dhe(Y, c)]$ the row corresponding to that constraint must be eliminated from $[Dh(Y, c)]$. But in addition, there must now be at least one $c_j > 0$, as we have specified the problem with the aggregate budget constraint always binding and aggregate production is now positive, so we
must also delete the row for the constraint \( c_j \). The matrix \([Dh(Y, c)]\) had a rank of \( 2I \); the question is, what will be the rank of the matrix \([DEh(Y, c)]\) formed by removing these two rows? When the row corresponding to \( Y_i \) is removed, the column corresponding to the derivative of \( h(Y, c) \) with respect to \( Y_i \) will have zeroes in every row below the first row. The first row contains the term \( 1 - G_Y' \), and if this term is zero we are left with a column of zeroes, which means that the column is redundant and the rank of the matrix is reduced to \( 2I - 1 \). In fact \( 1 - G_Y' \) will be zero, at a stationary point, so this column will be a column of zeroes. When the row corresponding to \( c_j \) is removed, we will be left with a column containing \( N_j > 0 \) in the first row and zeroes in every row below. This column is linearly independent of the other columns, and therefore the deletion of the row relating to \( c_j \) has not led to further reduction in the rank of the matrix. The end result then is that rank\( ([DEh(Y, c)]) = 2I - 1 \), which is the same as the number of effective constraints. Therefore the constraint qualification holds. If we now make the constraint on another output variable ineffective, the rank is reduced by 1, but so too is the number of effective constraints. So too if we make the constraint on another per capita consumption variable ineffective. In conclusion, for anything other than the degenerate case where all outputs and consumption levels are zero, we are assured that the constraint qualification is met for any feasible configuration of outputs and consumption levels.

**Critical points**

\[
\mathcal{L} = \sum_{i=1}^{I} N_i U(c_i) - \lambda \sum_{i=1}^{I} [Y_i - G(N_i, \theta_i, Y_i) - N_i c_i] + \sum_{i=1}^{I} \mu_i Y_i + \sum_{i=1}^{I} \phi_i c_i,
\]

The first-order conditions associated with the Lagrangian are:

\[
Y_i : \quad \lambda [1 - G_Y(N_i, \theta_i, Y_i)] + \mu_i = 0, \quad i = 1, \ldots, I,
\]

\[
c_i : \quad U''(c_i) - \lambda + \phi_i = 0, \quad i = 1, \ldots, I,
\]

\[
KT - Y : \quad Y_i \geq 0, \quad \mu_i \geq 0, \quad \mu_i Y_i = 0
\]

and

\[
KT - c : \quad c_i \geq 0, \quad \phi_i \geq 0, \quad \phi_i c_i = 0.
\]

Because we exclude the case where all \( Y_i = 0 \), aggregate output is greater than zero, and there must be some \( c_j > 0 \). Therefore \( \phi_j = 0 \) and \( U''(c_j) = \lambda \). Now suppose there is some \( c_k = 0 \). Then \( U'(0) - \lambda + \phi_k = 0 \) and \( U''(0) - U''(c_j) + \phi_k = 0 \). Because \( U'' < 0 \), we know that \( U''(0) - U''(c_j) > 0 \), and therefore \( \phi_k < 0 \). But this violates \( \phi_k \geq 0 \). Therefore we cannot have \( c_k = 0 \) for any region. We conclude therefore that \( c_i > 0 \) and \( \phi_i = 0 \) for all \( i \). We can therefore rewrite the first-order condition for the \( c_i \)'s as

\[
c_i : \quad U''(c_i) - \lambda = 0, \quad i = 1, \ldots, I.
\]
The matrix of second derivatives for the Lagrangian is

\[ [D^2 \mathcal{L}] = \begin{bmatrix}
-\lambda G''_1 & \ldots & 0 & 0 & \ldots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & \ldots & -\lambda G''_I & 0 & \ldots & 0 \\
0 & \ldots & 0 & U''_1 & \ldots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & \ldots & 0 & 0 & \ldots & U''_I
\end{bmatrix}, \]

where \( G''_i = G_{YY}(N_i, \theta_i, Y_i) \) and \( U''_i = U''(c_i) \). This matrix is negative definite everywhere, since it has only diagonal terms that are strictly negative. Either they are of the form \(-\lambda G''_i\), and negative because \( \lambda > 0 \) and \( G_{YY} > 0 \) by Proposition 1, or they are of the form \( U''_i \) and thus negative because \( U'' < 0 \). By virtue of negative definiteness, the maximum must be a global maximum.

E Decentralised output choice under complete information

E.1 Linear transfer schedule

Here we discuss the existence of a solution to this problem. I argue that it is not possible to appeal to the Weierstrass Theorem as a guarantee of the existence of a maximum. I then demonstrate that the constraint qualification for the Kuhn-Tucker method holds almost everywhere. This means that the Kuhn-Tucker conditions must be met, and I derive those conditions and demonstrate that there is a unique solution to them. Finally, I show that this unique solution is a maximiser.

Existence of a maximum

We have \( \{(T_i, t_i) : i = 1, \ldots, I\} \in \mathbb{R}^{2I} \), which is compact but not bounded. Moreover, it is not straightforward to identify useful bounds from the constraints of the problem. That being so, we will not resort to Weierstrass Theorem. Instead, we will proceed by considering the necessary conditions for a maximum, then we will use concavity to demonstrate that any maximum must be unique, and finally we will use the derivatives of the Lagrangian to show that there must indeed be a maximum.
Constraint qualification

Let $T = (T_1, \ldots, T_I)'$ and $t = (t_1, \ldots, t_I)'$. Form the constraint vector

$$h(T, t) = \begin{bmatrix} -\sum_{i=1}^{I} (T_i + t_i Y_i^*) \\ c_1'(T, t) \\ \vdots \\ c_I'(T, t) \end{bmatrix}$$

which has dimension $I + 1 \times 1$. First differentiate the central budget constraint. Theorem 3 tells us that $dY_i^*/dT_i = 0$ (see 10), from which it follows that

$$\frac{d}{dT_i} \left[ -\sum_{i=1}^{I} (T_i + t_i Y_i^*) \right] = -1,$$

and differentiating the budget constraint with respect to the marginal transfer rate gives

$$\frac{d}{dt_i} \left[ -\sum_{i=1}^{I} (T_i + t_i Y_i^*) \right] = -Y_i^* - t_i \frac{dY_i^*}{dt_i}.$$

The derivatives of optimised consumption can be had from (??) and (??). Thus the derivative of the constraint vector is

$$\begin{bmatrix} -1 & \ldots & -1 & -Y_1^* - t_1 \frac{dY_1^*}{dt_1} & \ldots & -Y_I^* - t_I \frac{dY_I^*}{dt_I} \\ \frac{1}{N_1} & \ldots & 0 & \frac{Y_1^* + t_1 \frac{dY_1^*}{dt_1}}{N_1} & \ldots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & \frac{1}{N_I} & 0 & 0 & \frac{Y_I^* + t_I \frac{dY_I^*}{dt_I}}{N_I} \end{bmatrix};$$

where the off-diagonal elements in the bottom two quadrants are all zeroes. This matrix has dimension $(I + 1) \times 2I$.

The first row of $[Dh(T, t)]$ is a linear combination of the last $I$ rows, but those last rows are independent of each other, so the rank of the matrix is $I$. This rank is less than the number of effective constraints, $I + 1$, so in the case where all the non-negativity constraints for optimised consumption are binding, the constraint qualification is not met. However, this is really a degenerate case, the case where the best achievable outcome entails zero consumption levels in every region. Now consider the case where the constraint for some $c_j^*$ is not binding. We form the the derivative matrix for the effective constraints $[Dh_{E}(T, t)]$ by deleting the row relating to $c_j$. Now, the first row is independent of the last $I - 1$ rows, which also are independent of each other. Therefore the rank of this matrix is $I$, which is the same as the number of effective constraints, and the constraint qualification holds. If we now set another consumption constraint non-binding, the rank of the
matrix falls to \( I - 1 \) and the number of effective constraints is \( I - 1 \) so the constraint qualification continues to hold, and so on.

In conclusion, the constraint qualification will always hold for this problem, except in the degenerate case where optimal consumption levels are zero in every region, a result that cannot arise so long as it is possible to produce a positive net output in at least one region.

**Critical points**

The derivative of the Lagrangian (27) with respect to the fixed elements \( T_i \) is

\[
\frac{dL}{dT_i} = \sum_{j=1}^{I} N_j U'(c_j^*) \frac{dc_j^*}{dT_i} - \mu + \sum_{j=1}^{I} \phi_j U'(c_j^*) \frac{dc_j^*}{dT_i}
= U'(c_i^*) - \mu + \frac{\phi_i}{N_i} U'(c_i^*), \quad i = 1, \ldots, I,
\]

where the second equality draws on (??) and (??). Thus the first-order condition with respect to \( T_i \) is

\[
T_i : \quad \frac{N_i}{N_i} U'(c_i^*) - \mu = 0, \quad i = 1, \ldots, I.
\]

And the derivative of the Lagrangian with respect to the marginal transfer rate \( t_i \) is

\[
\frac{dL}{dt_i} = \sum_{j=1}^{I} N_j U'(c_j^*) \frac{dc_j^*}{dt_i} - \mu \left( Y_i^* + \sum_{j=1}^{I} t_j \frac{dY_j^*}{dt_i} \right) + \sum_{j=1}^{I} \phi_j U'(c_j^*) \frac{dc_j^*}{dt_i}
= Y_i^* U'(c_i^*) - \mu \left( Y_i^* + t_i \frac{dY_i^*}{dt_i} \right) + \frac{\phi_i Y_i^*}{N_i} U'(c_i^*),
\]

where the second equality draws on (??) and (??). So, with minor rearrangement, the first-order condition with respect to \( t_i \) is

\[
t_i : \quad \frac{N_i}{N_i} Y_i^* U'(c_i^*) - \mu \left( Y_i^* + t_i \frac{dY_i^*}{dt_i} \right), = 0, \quad i = 1, \ldots, I.
\]

The Kuhn-Tucker conditions are

\[
KT - \mu : \quad \mu \geq 0, \quad \sum_{i=1}^{I} (T_i + t_i Y_i^*) \geq 0, \quad \mu \sum_{i=1}^{I} (T_i + t_i Y_i^*) = 0,
\]

and

\[
KT - \phi_i : \quad \phi_i \geq 0, \quad c_i^*(T, t) \geq 0, \quad \phi_i c_i^*(T, t) = 0, \quad i = 1, \ldots, I.
\]
If there is some region $j$ for which $c_j > 0$, then we must have $U'(c^*_j) = \mu$. And if there is some other region $k$ for which $\phi_k > 0$, then $(N_k + \phi_k)/N_k > 1$ and we must therefore have $U'(c^*_k) < \mu$. Therefore $U'(c^*_k) < U'(c^*_j)$, and since $U'' < 0$ it must be that $c^*_k > c^*_j$. But this contradicts $c^*_j > 0$ and $c^*_k = 0$. Therefore we cannot have $\phi_k > 0$ and must have $\phi_k = 0$. Therefore we require

$$U'(c^*_i) = \mu, \quad i = 1, \ldots, I,$$

which can be true only if the transfer policy is such that optimal per capita consumption levels are the same across regions.

F Decentralising choice of effort under incomplete information

References


